Network Working Group   H. Song, Ed.
Internet-Draft   T. Zhou
Intended status: Informational   ZB. Li
Expires: January 3, 2019   Huawei
                                                         G. Fioccola
                                                       Telecom Italia
                                                              ZQ. Li
                                                         China Mobile
                                                   P. Martinez-Julia
                                                                 NICT
                                                        L. Ciavaglia
                                                                Nokia
                                                             A. Wang
                                                       China Telecom
                                                         July 2, 2018

                    Toward a Network Telemetry Framework
                          draft-song-ntf-02

Abstract

   This document suggests the necessity of an architectural framework
   for network telemetry in order to meet the current and future network
   operation requirements. The defining characteristics of network
   telemetry shows a clear distinction from the conventional network OAM
   concept; hence the network telemetry demands new techniques and
   protocols. This document clarifies the terminologies and classifies
   the categories and components of a network telemetry framework. The
   requirements, challenges, existing solutions, and future directions
   are discussed for each category. The network telemetry framework and
   the taxonomy help to set a common ground for the collection of
   related works and put future technique and standard developments into
   perspective.

Requirements Language

   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 RFC 2119 [RFC2119].

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 January 3, 2019.

Copyright Notice

Table of Contents

1. Motivation

   The advance of AI/ML technologies gives networks an unprecedented
   opportunity to realize network autonomy with closed control loops.
   An intent-driven autonomous network is the logical next step for
   network evolution following SDN, aiming to reduce (or even eliminate)
   human labor, make the most efficient use of network resources, and
   provide better services more aligned with customer requirements.
   Although we still have a long way to reach the ultimate goal, the
   journey has started nevertheless.

   The storage and computing technologies are already mature enough to
   be able to retain and process a huge amount of data and make real-
   time inference. Tools based on machine learning technologies and big
   data analytics are powerful in detecting and reacting on network
   faults, anomalies, and policy violations. In turn, the network
   policy updates for planning, intrusion prevention, optimization, and
   self-healing can be applied. Some tools can even predict future
   events based on historical data.

   However, the networks fail to keep pace with such data need. The
   current network architecture, protocol suite, and system design are
   not ready yet to provide enough quality data. In the remaining of
   this section, first we identify a few key network operation use cases
   that network operators need the most. These use cases are also the
   essential functions of the future autonomous networks. Next, we show
   why the current network OAM techniques and protocols are not
   sufficient to meet the requirements of these use cases. The
   discussion underlines the need of a new brood of techniques and
   protocols which we put under an umbrella term - network telemetry.

1.1. Use Cases

   All these use cases involves the data extracted from the network data
   plane and sometimes from the network control plane and management
   plane.

Intent and Policy Compliance: Network policies are the rules that
   constraint the services for network access, provide differentiate
   within a service, or enforce specific treatment on the traffic.
   For example, a service function chain is a policy that requires
   the selected flows to pass through a set of network functions in
   order. An intents is a high-level abstract policy which requires
   a complex translation and mapping process before being applied on
   networks. While a policy is enforced, the compliance needs to be
   verified and monitored continuously.

SLA Compliance: A Service-Level Agreement (SLA) defines the level of
   service a user expects from a network operator, which include the
   metrics for the service measurement and remedy/penalty procedures
   when the service level misses the agreement. Users need to check
   if they get the service as promised and network operators need to
   evaluate how they can deliver the services that can meet the SLA.

Root Cause Analysis: Network failure often involves a sequence of
   chained events and the source of the failure is not
   straightforward to identify, especially when the failure is
   sporadic. While machine learning or other data analytics
   technologies can be used for root cause analysis, it up to the
   network to provide all the relevant data for analysis.

Load Balancing, Traffic Engineering, and Network Planning: Network
   operators are motivated to optimize their network utilization for
   better ROI or lower CAPEX, as well as differentiation across
   services and/or users of a given service. The first step is to
   know the real-time network conditions before applying policies to
   steer the user traffic or adjust the load balancing algorithm. In
   some cases network micro-bursts need to be detected in a very
   short time-frame so that fine grained traffic control can be
   applied to avoid possible network congestion. The long term
   network capacity planning and topology augmentation also rely on
   the accumulated data of the network operation.

Event Tracking and Prediction: Network visibility is critical for a
   healthy network operation. Numerous network events are of
   interest to network operators. For example, Network operators
   always want to learn where and why packets are dropped for an
   application flow. They also want to be warned by some early signs
   that some component is going to fail so the proper fix or
   replacement can be made in time.

1.2. Challenges

The conventional OAM techniques, as described in [RFC7276], are not
sufficient to support the above use cases for the following reasons:

o Most use cases need to continuously monitor the network and
   dynamically refine the data collection in real-time and
   interactively. The poll-based low-frequency data collection is
   ill-suited for these applications. Streaming data directly pushed
   from the data source is preferred.

o Various data is needed from any place ranging from the packet
   processing engine to the QoS traffic manager. Traditional data
   plane devices cannot provide the necessary probes. An open and
   programmable data plane is therefore needed.

o Many application scenarios need to correlate data from multiple
   sources (e.g., from distributed nodes or from different network
   plane). A piecemeal solution is often lacking the capability to
   consolidate the data from multiple sources. The composition of a
   complete solution, as partly proposed by ARCA
   [I-D.pedro-nmrg-anticipated-adaptation], will be empowered and
   guided by a comprehensive framework.

o The passive measurement techniques can either consume too much
   network resources and render too much redundant data, or lead to
   inaccurate results. The active measurement techniques are
   indirect, and they can interfere with the user traffic. We need
   techniques that can collect direct and on-demand data from user
   traffic.

1.3. Glossary

Before further discussion, we list some key terminology and acronyms
used in this documents. We make an intended distinction between
network telemetry and network OAM.

AI: Artificial Intelligence. Use machine-learning based
   technologies to automate network operation.

BMP: BGP Monitoring Protocol

DNP: Dynamic Network Probe

DPI: Deep Packet Inspection

gNMI: gPRC Network Management Interface

gRPC: gRPC Remote Procedure Call

IDN: Intent-Driven Network

IPFIX: IP Flow Information Export Protocol

IPFPM: IP Flow Performance Measurement

IOAM: In-situ OAM

NETCONF: Network Configuration Protocol

Network Telemetry: A general term for a new brood of network
   visibility techniques and protocols, with the characteristics
   defined in this document. Network telemetry enables smooth
   evolution toward intent-driven autonomous networks.

NMS: Network Management System

OAM: Operations, Administration, and Maintenance. A group of
   network management functions that provide network fault
   indication, fault localization, performance information, and data
   and diagnosis functions. Most conventional network monitoring
   techniques and protocols belong to network OAM.

SNMP: Simple Network Management Protocol

YANG: A data modeling language for NETCONF

YANG FSM: A YANG model to define device side finite state machine

YANG PUSH: A method to subscribe pushed data from remote YANG
   datastore

1.4. Network Telemetry

For a long time, network operators have relied upon protocols such as
SNMP [RFC1157] to monitor the network. SNMP can only provide limited
information about the network. Since SNMP is poll-based, it incurs
low data rate and high processing overhead. Such drawbacks make SNMP
unsuitable for today's automatic network applications.

Network telemetry has emerged as a mainstream technical term to refer
to the newer techniques of data collection and consumption,
distinguishing itself form the convention techniques for network OAM.
It is expected that network telemetry can provide the necessary
network visibility for autonomous networks, address the shortcomings

of conventional OAM techniques, and allow for the emergence of new
techniques bearing certain characterisitcs.

One key difference between the network telemetry and the network OAM
is that the network telemetry assumes an intelligent machine in the
center of a closed control loop, while the network OAM assumes the
human network operators in the middle of an open control loop. The
network telemetry can directly trigger the automated network
operation; The conventional OAM tools only help human operators to
monitor and diagnose the networks and guide manual network
operations. The different assumptions lead to very different
techniques.

Although the network telemetry techniques are just emerging and
subject to continuous evolution, several defining characteristics of
network telemetry have been well accepted:

o Push and Streaming: Instead of polling data from network devices,
   the telemetry collector subscribes to the streaming data pushed
   from the data source in network devices.

o Volume and Velocity: The telemetry data is intended to be consumed
   by machine rather than by human. Therefore, the data volume is
   huge and the processing is often in realtime.

o Normalization and Unification: Telemetry aims to address the
   overall network automation needs. The piecemeal solutions offered
   by the conventional OAM approach are no longer suitable. Efforts
   need to be made to normalize the data representation and unify the
   protocols.

o Model-based: The data is model-based which allows applications to
   configure and consume data with ease.

o Data Fusion: The data for a single application can come from
   multiple data sources (e.g., cross domain, cross device, and cross
   layer) and needs to be correlated to take effect.

o Dynamic and Interactive: Since the network telemetry means to be
   used in a closed control loop for network automation, it needs to
   run continuously and adapt to the dynamic and interactive queries
   from the network operation controller.

In addition, the ideal network telemetry solution should also support
the following features:

o In-Network Customization: The data can be customized in network at
   run-time to cater to the specific need of applications. This

needs the support of a programmable data plane which allows probes
to be deployed at flexible locations.

o Direct Data Plane Export: The data originated from data plane can
   be directly exported to the data consumer for efficiency,
   especially when the data bandwidth is large and the real-time
   processing is required.

o In-band Data Collection: In addition to the passive and active
   data collection approaches, the new hybrid approach allows to
   directly collect data for any target flow on its entire forwarding
   path.

o Non-intrusive: The telemetry system should not fall into the trap
   of the "observer effect". That is, it should not change the
   network behavior or affect the forwarding performance.

[2]. The Necessity of a Network Telemetry Framework

Big data analytics and machine-learning based AI technologies are
applied for network operation automation, relying on abundant data
from networks. The single-sourced and static data acquisition cannot
meet the data requirements. It is desirable to have a framework that
integrates multiple telemetry approaches from different layers, and
allows flexible combinations for different applications. The
framework will benefit application development for the following
reasons.

o The future autonomous networks will require a holistic view on
   network visibility. All the use cases and applications need to be
   supported uniformly and coherently under a single intelligent
   agent. Therefore, the protocols and mechanisms should be
   consolidated into a minimum yet comprehensive set. A telemetry
   framework can help to normalize the technique developments.

o Network visibility presents multiple viewpoints. For example, the
   device viewpoint takes the network infrastructure as the
   monitoring object from which the network topology and device
   status can be acquired; the traffic viewpoint takes the flows or
   packets as the monitoring object from which the traffic quality
   and path can be acquired. An application may need to switch its
   viewpoint during operation. It may also need to correlate a
   service and its network experience to acquire the comprehensive
   information.

o Applications require network telemetry to be elastic in order to
   efficiently use the network resource and reduce the performance
   impact. Routine network monitoring covers the entire network with

     low data sampling rate. When issues arise or trends emerge, the
     telemetry data source can be modified and the data rate can be
     boosted.

   o Efficient data fusion is critical for applications to reduce the
      overall quantity of data and improve the accuracy of analysis.

   So far, some telemetry related work has been done within IETF.
   However, this work is fragmented and scattered in different working
   groups. The lack of coherence makes it difficult to assemble a
   comprehensive network telemetry system and causes repetitive and
   redundant work.

   A formal network telemetry framework is needed for constructing a
   working system. The framework should cover the concepts and
   components from the standardization perspective. This document
   clarifies the layers on which the telemetry is exerted and decomposes
   the telemetry system into a set of distinct components that the
   existing and future work can easily map to.

3. Network Telemetry Framework

   Telemetry can be applied on the data plane, the control plane, and
   the management plane in a network, as well as other sources out of
   the network, as shown in Figure 1.

```
              +------------------------------+
              |                              |
              |       Network Operation      |<-------+
              |          Applications        |        |
              |                              |        |
              +------------------------------+        |
                  ^        ^          ^               |
                  |        |          |               |
                  V        |          V               V
              +-----------|---+-------------+  +-----------+
              |           |   |             |  |           |
              | Control Pl|ane|             |  | External  |
              | Telemetry | <--->           |  | Data and  |
              |           |   |             |  | Event     |
              |     ^     V   | Management  |  | Telemetry |
              +------|--------+  Plane      |  |           |
              |      V        | Telemetry   |  +-----------+
              |               |             |  |
              | Data Plane  <--->           |  |
              | Telemetry     |             |  |
              |               |             |  |
              +---------------+-------------+
```
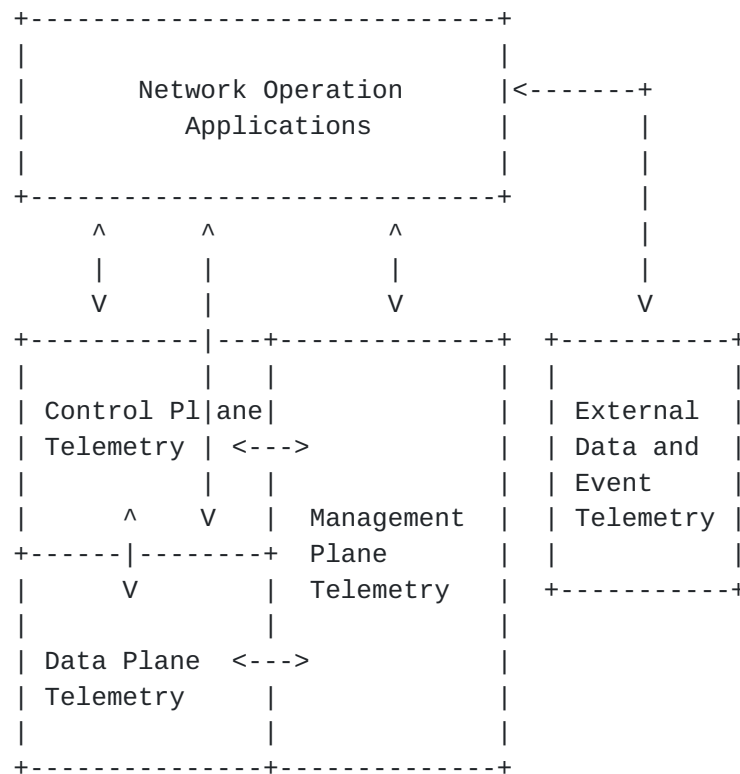
Figure 1: Layer Category of the Network Telemetry Framework

Note that the interaction with the network operation applications can be indirect. For example, in the management plane telemetry, the management plane may need to acquire data from the data plane. On the other hand, an application may involve more than one plane simultaneously. For example, an SLA compliance application may require both the data plane telemetry and the control plane telemetry.

At each plane, the telemetry can be further partitioned into five distinct components:

Data Source: Determine where the original data is acquired. The data source usually just provides raw data which needs further processing. A data source can be considered a probe. A probe can be statically installed or dynamically installed.

Data Subscription: Determine the protocol and channel for applications to acquire desired data. Data subscription is also responsible to define the desired data that might not be directly available form data sources. The subscription data can be described by a model. The model can be statically installed or dynamically installed.

Data Generation: The original data needs to be processed, encoded,
    and formatted in network devices to meet application subscription
    requirements. This may involve in-network computing and
    processing on either the fast path or the slow path in network
    devices.

Data Export: Determine how the ready data are delivered to
    applications.

Data Analysis and Storage: In this final step, data is consumed by
    applications or stored for future reference. Data analysis can be
    interactive. It may initiate further data subscription.

```
                +------------------------------+
                |                              |
                |     Data Analysis/Storage    |
                |                              |
                +------------------------------+
                     |                 ^
                     |                 |
                     V                 |
                +---------------+--------------+
                |               |              |
                | Data          | Data         |
                | Subscription  | Export       |
                |               |              |
                +---------------+--------------|
                |                              |
                |        Data Generation       |
                |                              |
                +------------------------------|
                |                              |
                |        Data Source           |
                |                              |
                +------------------------------+
```
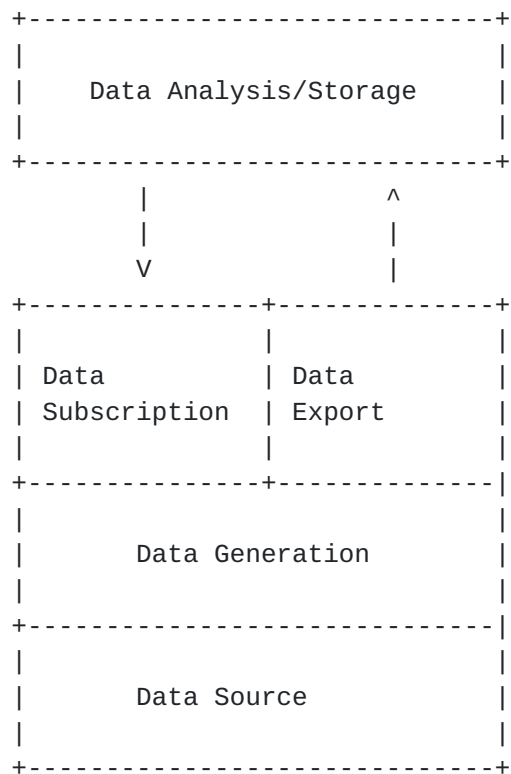
Figure 2: Components in the Network Telemetry Framework

Since most existing standard-related work belongs to the first four
components, in the remainder of the document, we focus on these
components only.

3.1. Existing Works Mapped in the Framework

The following table provides a non-exhaustive list of existing works
(mainly published in IETF and with the emphasis on the latest new
technologies) and shows their positions in the framework.

```
+----------+-------------+--------------+-------------+
|          | Management  | Control      | Data        |
|          | Plane       | Plane        | Plane       |
+----------+-------------+--------------+-------------+
|          | YANG Data   | Control Proto.| Flow/Packet |
| Data     | Store       | Network State | Statistics  |
| Source   |             |              | States      |
|          |             |              | DPI         |
+----------+-------------+--------------+-------------+
|          | gPRC        | NETCONF/YANG | NETCONF/YANG |
| Data     | YANG PUSH   | BGP          | YANG FSM    |
| Subscribe|             |              |             |
|          |             |              |             |
+----------+-------------+--------------+-------------+
|          | Soft DNP    | Soft DNP     | In-situ OAM |
| Data     |             |              | IPFPM       |
| Generation|            |              | Hard DNP    |
|          |             |              |             |
+----------+-------------+--------------+-------------+
|          | gRPC        | BMP          | IPFIX       |
| Data     | YANG PUSH   |              | UDP         |
| Export   | UDP         |              |             |
|          |             |              |             |
+----------+-------------+--------------+-------------+
```

Figure 3: Existing Work

3.2. Management Plane Telemetry

3.2.1. Requirements and Challenges

   The management plane of the network element interacts with the
   Network Management System (NMS), and provides information such as
   performance data, network logging data, network warning and defects
   data, and network statistics and state data. Some legacy protocols
   are widely used for the management plane, such as SNMP and Syslog,
   but these protocols do not meet the requirements of the automatic
   network operation applications.

   New management plane telemetry protocols should consider the
   following requirements:

   Convenient Data Subscription: An application should have the freedom
      to choose the data export means such as the data types and the
      export frequency.

Structured Data: For automatic network operation, machines will
   replace human for network data comprehension. The schema
   languages such as YANG can efficiently describe structured data
   and normalize data encoding and transformation.

High Speed Data Transport: In order to retain the information, a
   server needs to send a large amount of data at high frequency.
   Compact encoding formats are needed to compress the data and
   improve the data transport efficiency. The push mode, by
   replacing the poll mode, can also reduce the interactions between
   clients and servers, which help to improve the server's
   efficiency.

3.2.2. Push Extensions for NETCONF

NETCONF [RFC6241] is one popular network management protocol, which
is also recommended by IETF. Although it can be used for data
collection, NETCONF is good at configurations. YANG Push
[I-D.ietf-netconf-yang-push] extends NETCONF and enables subscriber
applications to request a continuous, customized stream of updates
from a YANG datastore. Providing such visibility into changes made
upon YANG configuration and operational objects enables new
capabilities based on the remote mirroring of configuration and
operational state. Moreover, distributed data collection mechanism
[I-D.zhou-netconf-multi-stream-originators] via UDP based publication
channel [I-D.ietf-netconf-udp-pub-channel] provides enhanced
efficiency for the NETCONF based telemetry.

3.2.3. gRPC Network Management Interface

gRPC Network Management Interface (gNMI)
[I-D.openconfig-rtgwg-gnmi-spec] is a network management protocol
based on the gRPC [I-D.kumar-rtgwg-grpc-protocol] RPC (Remote
Procedure Call) framework. With a single gRPC service definition,
both configuration and telemetry can be covered. gRPC is an HTTP/2
[RFC7540] based open source micro service communication framework.
It provides a number of capabilities that makes it well-suited for
network telemetry, including:

o Full-duplex streaming transport model combined with a binary
   encoding mechanism provided further improved telemetry efficiency.

o gRPC provides higher-level features consistency across platforms
   that common HTTP/2 libraries typically do not. This
   characteristic is especially valuable for the fact that telemetry
   data collectors normally reside on a large variety of platforms.

o The built-in load-balancing and failover mechanism.

3.3. Control Plane Telemetry

3.3.1. Requirements and Challenges

   The control plane telemetry refers to the health condition monitoring
   of different network protocols, which covers Layer 2 to Layer 7.
   Keeping track of the running status of these protocols is beneficial
   for detecting, localizing, and even predicting various network
   issues, as well as network optimization, in real-time and in fine
   granularities.

   One of the most challenging problems for the control plane telemetry
   is how to correlate the E2E Key Performance Indicators (KPI) to a
   specific layer's KPIs. For example, an IPTV user may describe his
   User Experience (UE) by the video fluency and definition. Then in
   case of an unusually poor UE KPI or a service disconnection, it is
   non-trivial work to delimit and localize the issue to the responsible
   protocol layer (e.g., the Transport Layer or the Network Layer), the
   responsible protocol (e.g., ISIS or BGP at the Network Layer), and
   finally the responsible device(s) with specific reasons.

   Traditional OAM-based approaches for control plane KPI measurement
   include PING (L3), Tracert (L3), Y.1731 (L2) and so on. One common
   issue behind these methods is that they only measure the KPIs instead
   of reflecting the actual running status of these protocols, making
   them less effective or efficient for control plane troubleshooting
   and network optimization. An example of the control plane telemetry
   is the BGP monitoring protocol (BMP), it is currently used to
   monitoring the BGP routes and enables rich applications, such as BGP
   peer analysis, AS analysis, prefix analysis, security analysis, and
   so on. However, the monitoring of other layers, protocols and the
   cross-layer, cross-protocol KPI correlations are still in their
   infancies (e.g., the IGP monitoring is missing), which require
   substantial further research.

3.3.2. BGP Monitoring Protocol

   BGP Monitoring Protocol (BMP) [RFC7854] is used to monitor BGP
   sessions and intended to provide a convenient interface for obtaining
   route views.

   The BGP routing information is collected from the monitored device(s)
   to the BMP monitoring station by setting up the BMP TCP session. The
   BGP peers are monitored by the BMP Peer Up and Peer Down
   Notifications. The BGP routes (including Adjacency_RIB_In [RFC7854],
   Adjacency_RIB_out [I-D.ietf-grow-bmp-adj-rib-out], and Local_Rib
   [I-D.ietf-grow-bmp-local-rib] are encapsulated in the BMP Route
   Monitoring Message and the BMP Route Mirroring Message, in the form

of both initial table dump and real-time route update. In addition,
BGP statistics are reported through the BMP Stats Report Message,
which could be either timer triggered or event driven. More BMP
extensions can be explored to enrich the applications of BGP
monitoring.

3.4. Data Plane Telemetry

3.4.1. Requirements and Challenges

An effective data plane telemetry system relies on the data that the
network device can expose. The data's quality, quantity, and
timeliness must meet some stringent requirements. This raises some
challenges to the network data plane devices where the first hand
data originate.

o A data plane device's main function is user traffic processing and
   forwarding. While supporting network visibility is important, the
   telemetry is just an auxiliary function and it should not impede
   normal traffic processing and forwarding (i.e., the performance is
   not lowered and the behavior is not altered due to the telemetry
   functions).

o The network operation applications requires end-to-end visibility
   from various sources, which results in a huge volume of data.
   However, the sheer data quantity should not stress the network
   bandwidth, regardless of the data delivery approach (i.e., through
   in-band or out-of-band channels).

o The data plane devices must provide the data in a timely manner
   with the minimum possible delay. Long processing, transport,
   storage, and analysis delay can impact the effectiveness of the
   control loop and even render the data useless.

o The data should be structured and labeled, and easy for
   applications to parse and consume. At the same time, the data
   types needed by applications can vary significantly. The data
   plane devices need to provide enough flexibility and
   programmability to support the precise data provision for
   applications.

o The data plane telemetry should support incremental deployment and
   work even though some devices are unaware of the system. This
   challenge is highly relevant to the standards and legacy networks.

The industry has agreed that the data plane programmability is
essential to support network telemetry. Newer data plane chips are

all equipped with advanced telemetry features and provide flexibility
to support customized telemetry functions.

3.4.2. Technique Classification

There can be multiple possible dimensions to classify the data plane
telemetry techniques.

Active and Passive: The active and passive methods (as well as the
   hybrid types) are well documented in [RFC7799]. The passive
   methods include TCPDUMP, IPFIX [RFC7011], sflow, and traffic
   mirror. These methods usually have low data coverage. The
   bandwidth cost is very high in oreder to improve the data
   coverage. On the other hand, the active methods include Ping,
   Traceroute, OWAMP [RFC4656], and TWAMP [RFC5357]. These methods
   are intrusive and only provide indirect network measurement
   results. The hybrid methods, including in-situ OAM
   [I-D.brockners-inband-oam-requirements], IPFPM [RFC8321], and
   Multipoint Alternate Marking
   [I-D.fioccola-ippm-multipoint-alt-mark], provide a well-balanced
   and more flexible approach. However, these methods are also more
   complex to implement.

In-Band and Out-of-Band: The telemetry data, before being exported
   to some collector, can be carried in user packets. Such methods
   are considered in-band (e.g., in-situ OAM
   [I-D.brockners-inband-oam-requirements]). If the telemetry data
   is directly exported to some collector without modifying the user
   packets, Such mothods are considered out-of-band (e.g., postcard-
   based INT). It is possible to have hybrid methods. For example,
   only the telemetry instruction or partitial data is carried by
   user packets (e.g., IPFPM [RFC8321]).

E2E and In-Network: Some E2E methods start from and end at the
   network end hosts (e.g., Ping). The other methods work in
   networks and are transparent to end hosts. However, if needed,
   the in-network methods can be easily extended into end hosts.

Flow, Path, and Node: Depending on the telemetry objective, the
   methods can be flow-based (e.g., in-situ OAM
   [I-D.brockners-inband-oam-requirements]), path-based (e.g.,
   Traceroute), and node-based (e.g., IPFIX [RFC7011]).

3.4.3. The IPFPM technology

The Alternate Marking method is efficient to perform packet loss,
delay, and jitter measurements both in an IP and Overlay Networks, as

presented in IPFPM [RFC8321] and
[I-D.fioccola-ippm-multipoint-alt-mark].

This technique can be applied to point-to-point and multipoint-to-multipoint flows. Alternate Marking creates batches of packets by alternating the value of 1 bit (or a label) of the packet header. These batches of packets are unambiguously recognized over the network and the comparison of packet counters for each batch allows the packet loss calculation. The same idea can be applied to delay measurement by selecting ad hoc packets with a marking bit dedicated for delay measurements.

Alternate Marking method needs two counters each marking period for each flow under monitor. For instance, by considering n measurement points and m monitored flows, the order of magnitude of the packet counters for each time interval is n*m*2 (1 per color).

Since networks offer rich sets of network performance measurement data (e.g packet counters), traditional approaches run into limitations. One reason is the fact that the bottleneck is the generation and export of the data and the amount of data that can be reasonably collected from the network. In addition, management tasks related to determining and configuring which data to generate lead to significant deployment challenges.

Multipoint Alternate Marking approach, described in
[I-D.fioccola-ippm-multipoint-alt-mark], aims to resolve this issue and makes the performance monitoring more flexible in case a detailed analysis is not needed.

An application orchestrates network performance measurements tasks across the network to allow an optimized monitoring and it can calibrate how deep can be obtained monitoring data from the network by configuring measurement points roughly or meticulously.

Using Alternate Marking, it is possible to monitor a Multipoint Network without examining in depth by using the Network Clustering (subnetworks that are portions of the entire network that preserve the same property of the entire network, called clusters). So in case there is packet loss or the delay is too high the filtering criteria could be specified more in order to perform a detailed analysis by using a different combination of clusters up to a per-flow measurement as described in IPFPM [RFC8321].

In summary, an application can configure initially an end to end monitoring between ingress points and egress points of the network. If the network does not experiment issues, this approximate monitoring is good enough and is very cheap in terms of network

resources. But, in case of problems, the application becomes aware
of the issues from this approximate monitoring and, in order to
localize the portion of the network that has issues, configures the
measurement points more exhaustively. So a new detailed monitoring
is performed. After the detection and resolution of the problem the
initial approximate monitoring can be used again.

3.4.4. Dynamic Network Probe

Hardware based Dynamic Network Probe (DNP) [I-D.song-opsawg-dnp4iq]
provides a programmable means to customize the data that an
application collects from the data plane. A direct benefit of DNP is
the reduction of the exported data. A full DNP solution covers
several components including data source, data subscription, and data
generation. The data subscription needs to define the custom data
which can be composed and derived from the raw data sources. The
data generation takes advantage of the moderate in-network computing
to produce the desired data.

While DNP can introduce unforeseeable flexibility to the data plane
telemetry, it also faces some challenges. It requires a flexible
data plane that can be dynamically reprogrammed at run-time. The
programming API is yet to be defined.

3.4.5. IP Flow Information Export (IPFIX) protocol

Traffic on a network can be seen as a set of flows passing through
network elements. IP Flow Information Export (IPFIX) [RFC7011]
provides a means of transmitting traffic flow information for
administrative or other purposes. A typical IPFIX enabled system
includes a pool of Metering Processes collects data packets at one or
more Observation Points, optionally filters them and aggregates
information about these packets. An Exporter then gathers each of
the Observation Points together into an Observation Domain and sends
this information via the IPFIX protocol to a Collector.

3.4.6. In-Situ OAM

Traditional passive and active monitoring and measurement techniques
are either inaccurate or resource-consuming. It is preferable to
directly acquire data associated with a flow's packets when the
packets pass through a network. In-situ OAM (iOAM)
[I-D.brockners-inband-oam-requirements], a data generation technique,
embeds a new instruction header to user packets and the instruction
directs the network nodes to add the requested data to the packets.
Thus, at the path end the packet's experience on the entire
forwarding path can be collected. Such firsthand data is invaluable
to many network OAM applications.

However, iOAM also faces some challenges. The issues on performance
impact, security, scalability and overhead limits, encapsulation
difficulties in some protocols, and cross-domain deployment need to
be addressed.

3.5. External Data and Event Telemetry

Events that occur outside the boundaries of the network system are
another important source of telemetry information. Correlating both
internal telemetry data and external events with the requirements of
network systems, as presented in Exploiting External Event Detectors
to Anticipate Resource Requirements for the Elastic Adaptation of
SDN/NFV Systems [I-D.pedro-nmrg-anticipated-adaptation], provides a
strategic and functional advantage to management operations.

3.5.1. Requirements and Challenges

As with other sources of telemetry information, the data and events
must meet strict requirements, especially in terms of timeliness,
which is essential to properly incorporate external event information
to management cycles. Thus, the specific challenges are described as
follows:

o The role of external event detector can be played by multiple
   elements, including hardware (e.g. physical sensors, such as
   seismometers) and software (e.g. Big Data sources that analyze
   streams of information, such as Twitter messages). Thus, the
   transmitted data must support different shapes but, at the same
   time, follow a common but extensible ontology.

o Since the main function of the external event detectors is
   actually to perform the notifications, their timeliness is
   assumed. However, once messages have been dispatched, they must
   be quickly collected and inserted into the control plane with
   variable priority, which will be high for important sources and/or
   important events and low for secondary ones.

o The ontology used by external detectors must be easily adopted by
   current and future devices and applications. Therefore, it must
   be easily mapped to current information models, such as in terms
   of YANG.

Organizing together both internal and external telemetry information
will be key for the general exploitation of the management
possibilities of current and future network systems, as reflected in
the incorporation of cognitive capabilities to new hardware and
software (virtual) elements.

4. Security Considerations

   TBD

5. IANA Considerations

   This document includes no request to IANA.

6. Contributors

   The other main contributors of this document are listed as follows.

   o James N. Guichard, Huawei

   o Yunan Gu, Huawei

7. Acknowledgments

   We would like to thank Victor Liu and others who have provided
   helpful comments and suggestions to improve this document.

8. References

8.1. Normative References

   [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>.

8.2. Informative References

   [I-D.brockners-inband-oam-requirements]
              Brockners, F., Bhandari, S., Dara, S., Pignataro, C.,
              Gredler, H., Leddy, J., Youell, S., Mozes, D., Mizrahi,
              T., <>, P., and r. remy@barefootnetworks.com,
              "Requirements for In-situ OAM", draft-brockners-inband-
              oam-requirements-03 (work in progress), March 2017.

   [I-D.fioccola-ippm-multipoint-alt-mark]
              Fioccola, G., Cociglio, M., Sapio, A., and R. Sisto,
              "Multipoint Alternate Marking method for passive and
              hybrid performance monitoring", draft-fioccola-ippm-
              multipoint-alt-mark-04 (work in progress), June 2018.

[I-D.ietf-grow-bmp-adj-rib-out]
          Evens, T., Bayraktar, S., Lucente, P., Mi, K., and S.
          Zhuang, "Support for Adj-RIB-Out in BGP Monitoring
          Protocol (BMP)", draft-ietf-grow-bmp-adj-rib-out-01 (work
          in progress), March 2018.

[I-D.ietf-grow-bmp-local-rib]
          Evens, T., Bayraktar, S., Bhardwaj, M., and P. Lucente,
          "Support for Local RIB in BGP Monitoring Protocol (BMP)",
          draft-ietf-grow-bmp-local-rib-01 (work in progress),
          February 2018.

[I-D.ietf-netconf-udp-pub-channel]
          Zheng, G., Zhou, T., and A. Clemm, "UDP based Publication
          Channel for Streaming Telemetry", draft-ietf-netconf-udp-
          pub-channel-03 (work in progress), July 2018.

[I-D.ietf-netconf-yang-push]
          Clemm, A., Voit, E., Prieto, A., Tripathy, A., Nilsen-
          Nygaard, E., Bierman, A., and B. Lengyel, "YANG Datastore
          Subscription", draft-ietf-netconf-yang-push-17 (work in
          progress), July 2018.

[I-D.kumar-rtgwg-grpc-protocol]
          Kumar, A., Kolhe, J., Ghemawat, S., and L. Ryan, "gRPC
          Protocol", draft-kumar-rtgwg-grpc-protocol-00 (work in
          progress), July 2016.

[I-D.openconfig-rtgwg-gnmi-spec]
          Shakir, R., Shaikh, A., Borman, P., Hines, M., Lebsack,
          C., and C. Morrow, "gRPC Network Management Interface
          (gNMI)", draft-openconfig-rtgwg-gnmi-spec-01 (work in
          progress), March 2018.

[I-D.pedro-nmrg-anticipated-adaptation]
          Martinez-Julia, P., "Exploiting External Event Detectors
          to Anticipate Resource Requirements for the Elastic
          Adaptation of SDN/NFV Systems", draft-pedro-nmrg-
          anticipated-adaptation-02 (work in progress), June 2018.

[I-D.song-opsawg-dnp4iq]
          Song, H. and J. Gong, "Requirements for Interactive Query
          with Dynamic Network Probes", draft-song-opsawg-dnp4iq-01
          (work in progress), June 2017.

[I-D.zhou-netconf-multi-stream-originators]
          Zhou, T., Zheng, G., Voit, E., Clemm, A., and A. Bierman,
          "Subscription to Multiple Stream Originators", draft-zhou-
          netconf-multi-stream-originators-02 (work in progress),
          May 2018.

[RFC1157] Case, J., Fedor, M., Schoffstall, M., and J. Davin,
          "Simple Network Management Protocol (SNMP)", RFC 1157,
          DOI 10.17487/RFC1157, May 1990,
          <https://www.rfc-editor.org/info/rfc1157>.

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M.
          Zekauskas, "A One-way Active Measurement Protocol
          (OWAMP)", RFC 4656, DOI 10.17487/RFC4656, September 2006,
          <https://www.rfc-editor.org/info/rfc4656>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J.
          Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)",
          RFC 5357, DOI 10.17487/RFC5357, October 2008,
          <https://www.rfc-editor.org/info/rfc5357>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed.,
          and A. Bierman, Ed., "Network Configuration Protocol
          (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011,
          <https://www.rfc-editor.org/info/rfc6241>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken,
          "Specification of the IP Flow Information Export (IPFIX)
          Protocol for the Exchange of Flow Information", STD 77,
          RFC 7011, DOI 10.17487/RFC7011, September 2013,
          <https://www.rfc-editor.org/info/rfc7011>.

[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>.

[RFC7540] Belshe, M., Peon, R., and M. Thomson, Ed., "Hypertext
          Transfer Protocol Version 2 (HTTP/2)", RFC 7540,
          DOI 10.17487/RFC7540, May 2015,
          <https://www.rfc-editor.org/info/rfc7540>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with
          Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799,
          May 2016, <https://www.rfc-editor.org/info/rfc7799>.

   [RFC7854]   Scudder, J., Ed., Fernando, R., and S. Stuart, "BGP
               Monitoring Protocol (BMP)", RFC 7854,
               DOI 10.17487/RFC7854, June 2016,
               <https://www.rfc-editor.org/info/rfc7854>.

   [RFC8321]   Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli,
               L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi,
               "Alternate-Marking Method for Passive and Hybrid
               Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321,
               January 2018, <https://www.rfc-editor.org/info/rfc8321>.

Authors' Addresses

   Haoyu Song (editor)
   Huawei
   2330 Central Expressway
   Santa Clara
   USA

   Email: haoyu.song@huawei.com

   Tianran Zhou
   Huawei
   156 Beiqing Road
   Beijing, 100095
   P.R. China

   Email: zhoutianran@huawei.com

   Zhenbin Li
   Huawei
   156 Beiqing Road
   Beijing, 100095
   P.R. China

   Email: lizhenbin@huawei.com

   Giuseppe Fioccola
   Telecom Italia
   Via Reiss Romoli, 274
   Torino 10148
   Italy

   Email: giuseppe.fioccola@telecomitalia.it

Zhenqiang Li
China Mobile
No. 32 Xuanwumenxi Ave., Xicheng District
Beijing, 100032
P.R. China

Email: lizhenqiang@chinamobile.com

Pedro Martinez-Julia
NICT
4-2-1, Nukui-Kitamachi
Koganei, Tokyo 184-8795
Japan

Phone: +81 42 327 7293
Email: pedro@nict.go.jp

Laurent Ciavaglia
Nokia
Villarceaux 91460
France

Email: laurent.ciavaglia@nokia.com

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing, 102209
P.R. China

Email: wangaj.bri@chinatelecom.cn