Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 2, 2018

H. Song, Ed.
T. Zhou
Z. Li
Huawei

March 1, 2018

# Toward a Network Telemetry Framework
# draft-song-ntf-00

## Abstract

This document suggests the necessity of a framework of network telemetry and articulates the categories and components of such a framework. The requirement, challenges, existing solutions, and future directions are discussed for each category of the framework. The framework for network telemetry helps to set a common ground for the collection of related works and put future 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].

## Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 2, 2018.

## Copyright Notice

Table of Contents



## 1. Motivation

Intent-based automatic network is the logical next step of network evolution, aiming to reduce human labor, make the most efficient use of network resources, and provide better services. Tools based on machine learning technologies and big data analytics are powerful for faults, anomaly, pattern, and policy violation detection. Some tools can even predict future events based on history data. The

observation and inference from the network data can guide the network policy updates for planning, intrusion prevention, optimization, and self-healing. A closed control loop is therefore achieved.

Network OAM provides necessary visibilities to a network. It plays an important role in Intend-based Networks (IBN).

## 1.1. Use Cases

Specifically, we have identified a few key network OAM use cases that service providers need the most. All these use cases involves the data extracted from the network data plane and sometimes from the network control plane and management plane:

Policy Compliance: Network policies are the rules that constraint the services for network access. For example, a service function chain is a policy that requires the selected flows to pass through a set of network functions in order. While a policy is enforced, the compliance needs to be monitored continuously.

SLA Compliance: A service-level agreement defines the level of service a user expects from a service provider, 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 service providers need to evaluate how they can deliver the services that can meet the Service Level Agreement (SLA).

Root Cause Analysis: Network failure often involves a sequence of chain events and the source of the failure is not straightforward to identify, especially when t:0he 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 and Traffic Engineering: Service providers are motivated to optimize their network utilization for better ROI or lower CAPEX. The first step is to know the real-time network condition before applying policies to steer the user traffic or adjust the load balancing algorithm. In some cases the network micro-bursts need to be detected in a very short time-frame so does the fine grained traffic control can be applied to avoid the possible network congestion.

Packet Drop Detection: Sporadic packet drops in networks are notoriously hard to locate and debug. Network operators are plagued by the lack of tools that can identify the packet drop

locations and reasons. Both active and passive measurements are not very effective in solving this problem.

These use cases show that the conventional OAM techniques are not enough for the following reasons:

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

- Various data are 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.

- 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 lack of the capability to consolidate the data from multiple sources. The composition of a complete solution can be guided by a comprehensive framework.

- 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.

## 1.2. Terminology and Abbreviations

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

BMP: BGP Monitoring Protocol

DNP: Dynamic Network Probe

gNMI: gPRC Network Managment Interface

gRPC: gRPC Remote Procesure Call

IBN: Intent-Based 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 techniques to gain network visibility, through network data collection for analysis and measurement.

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.

SNMP: Simple Network Managment 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

## 1.3. Network Telemetry

For a long time, network OAM applications rely on protocols such as SNMP [RFC1157] to monitor the networks. 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 technologies of data collection and consumption in the IBN paradigm, distinguishing itself form the convention technologies for network OAM. It is expected that the network telemetry can provide the necessary network visibility for automated network OAM, address the shortcomings of the conventional technologies, and allow the emergence of new technologies.

Although the network telemetry technologies continue evolving, several defining characteristics of network telemetry have been well accepted:

- Instead of polling data from the network devices, the telemetry collector subscribes the streaming data pushed from the data source in network devices.

- The data is normalized and encoded efficiently for export.
- The data is model-based which allows applications to configure and consume data with ease.

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

- The data can be customized at runtime to cater the specific need of applications. This needs the support of a programmable data plane which allows probes to be deployed at flexible locations.
- The data for a single application can come from multiple data sources (e.g., cross domain, cross device, and cross layer) and need to be correlated to take effect.

### 1.4. The Necessity of a Network Telemetry Framework

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

- 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.
- Applications require the network telemetry to be elastic in order to efficiently use the network resource and reduce the performance impact. The routine network monitoring covers the entire network with low data sampling rate. When issues arise or trends emerge, the telemetry data source can be refocused and the data rate can be boosted.
- 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 works have been done within IETF. However, these works are 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 works.

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 works can easily map to.

## 2. Network Telemetry Framework

The telemetry can be applied on the data plane, the control plane, and the management plane in a network, as shown in Figure 1.

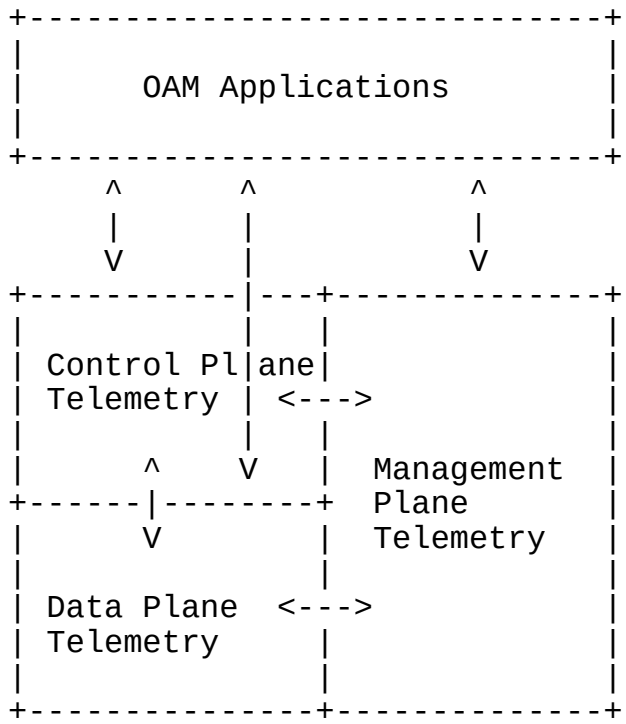


Figure 1: Layer Category of the Network Telemetry Framework

Note that the interaction with OAM 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 OAM 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 provide 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 directly available form data sources. The subscribe 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: In this final step, data is consumed by applications. Data analysis can be interactive. It may initiate further data subscription.

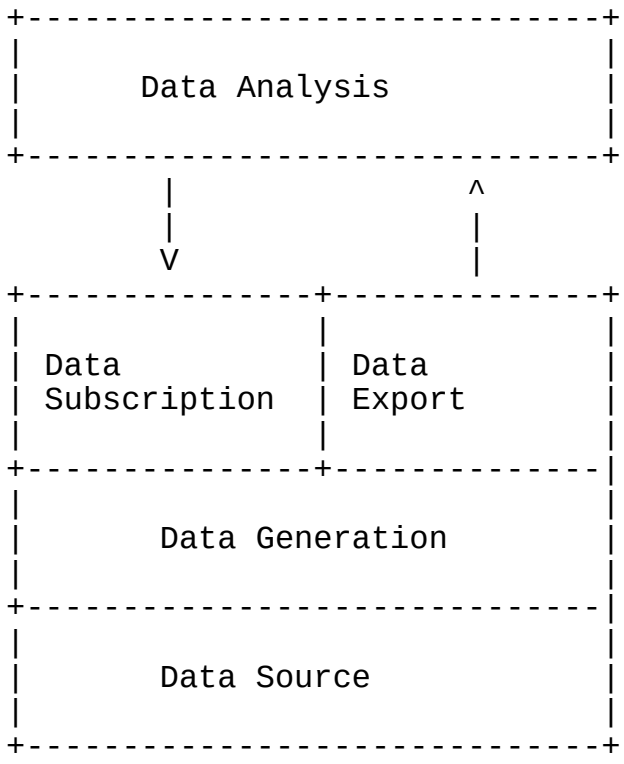


Figure 2: Components in the Network Telemetry Framework

Since most existing standard-related works belong to the first four components, in the remaining of the document, we focus on these components only.

## 2.1. 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.

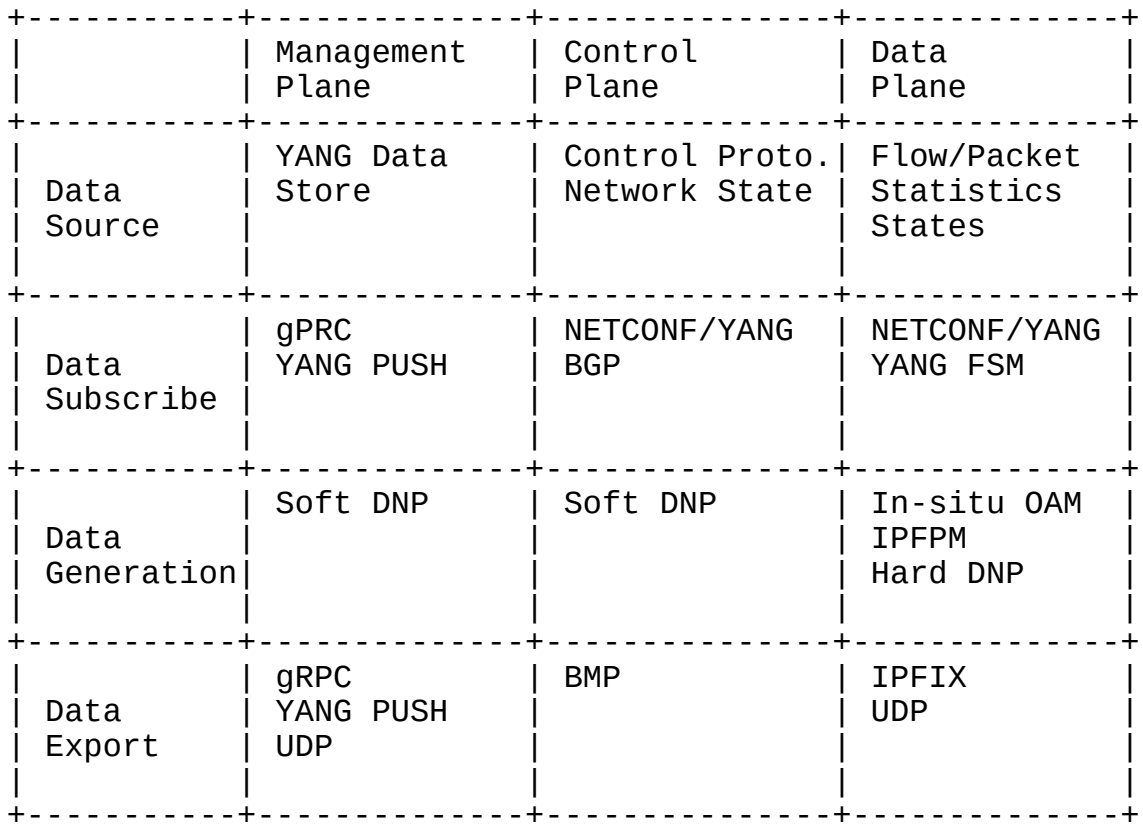


Figure 3: Existing Works

## 2.2. Management Plane Telemetry

### 2.2.1. 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 OAM 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 OAM, machine 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 need to send a large amount of data at high frequency. Compact encoding format is 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.

### 2.2.2. 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.

### 2.2.3. 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:

- Full-duplex streaming transporting model combined with a binary encoding mechanism provided further improved telemetry efficiency.
- 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 are normally resides on a large variety of platforms.
- The build in load balancing and failover mechanism.

## 2.3. Control Plane Telemetry

### 2.3.1. Requirements and Challenges

The control plane runs the routing protocol (e.g., BGP, OSPF, and IS-IS) to calculate the routing table for a network device. The control plane telemetry monitors the routing protocols to ensure they are working properly.

### 2.3.2. 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 data is collected from the Adjacency-RIB-In routing tables, which are the pre-policy tables, meaning that the routes in these tables have not been filtered or modified by routing policies. So the monitoring station can receive all routes, not just the active routes.

## 2.4. Data Plane Telemetry

### 2.4.1. 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.

- 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).

- The network OAM 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).

- 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.

- 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.

- 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.

#### 2.4.2. 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 runtime. The programming API is yet to be defined.

#### 2.4.3. 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.

#### 2.4.4. 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. Security Considerations

TBD

## 4. IANA Considerations

This document includes no request to IANA.

## 5. Contributors

The other contributors of this document are listed as follows.

o Yunan Gu, Huawei

## 6. Acknowledgments

TBD.

## 7. References

### 7.1. Normative References

[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>.

[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>.

[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>.

[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>.

[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>.

## 7.2. 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.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-01 (work in progress), November 2017.

[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-15 (work in progress), February 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-00 (work in progress), March 2017.

[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-01 (work in progress), November 2017.

## 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