

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 23, 2021

Y. Gu
S. Chen
Huawei
Y. Qu
Futurewei
H. Chen
China Telecom
Z. Li
Huawei
February 19, 2021

Network Monitoring For IGP
draft-gu-opsawg-network-monitoring-igp-01

Abstract

To evolve towards automated network OAM (Operations, administration and management), the monitoring of control plane protocols is a fundamental necessity. This document proposes network monitoring for IGP to facilitate troubleshooting by collecting the IGP monitoring data and reporting it to the network monitoring server in real-time. In this document, the operations of network monitoring for ISIS are described, and the corresponding network monitoring message types and message formats are defined.

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 August 23, 2021.

Internet-Draft

Network Monitoring For IGP

February 2021

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	Motivation	2
1.2.	Overview	4
2.	Terminology	5
3.	Use Cases	5
3.1.	IS-IS Route Flapping	5
3.2.	IS-IS LSDB Synchronization Failure	6
4.	Message Format	7
4.1.	Protocol Selection Options	7
4.2.	Message Types	7
4.3.	Message Format	8
4.3.1.	Common Header	8
4.3.2.	Per Adjacency Header	8
4.3.3.	Initiation Message	9
4.3.4.	Adjacency Status Change Notification	9
4.3.5.	ISIS Statistic Report Message	11
4.3.6.	IS-IS PDU Monitoring Message	12
4.3.7.	Termination Message	13
5.	IANA	13
6.	Contributors	13
7.	Acknowledgments	13
8.	References	13
	Authors' Addresses	14

[1.](#) Introduction

[1.1.](#) Motivation

The requirement for better network OAM approaches has been greatly driven by the network evolvement. The concept of network Telemetry has been proposed to meet the current and future OAM demands w.r.t.,

massive and real-time data storage, collection, process, export, and analysis, and an architectural framework of existing Telemetry approaches is introduced in [[I-D.song-ntf](#)]. Network Telemetry provides visibility to the network health conditions, and is beneficial for faster network troubleshooting, network OpEx (operating expenditure) reduction, and network optimization. Telemetry can be applied to the data plane, control plane and management plane. There have been various methods proposed for each plane:

- o Management plane: For example, SNMP (Simple Network Management Protocol) [[RFC1157](#)], NETCONF (Network Configuration Protocol) [[RFC6241](#)] and gNMI (gRPC Network Management Interface) [[I-D.openconfig-rtgwg-gnmi-spec](#)] are three typical widely adopted management plane Telemetry approaches. Various YANG modules are defined for network operational state retrieval and configuration management. Subscription to specific YANG datastore can be realized in combination with gRPC/NETCONF.
- o Data plane: For example, In-situ OAM (iOAM) [[I-D.brockners-inband-oam-requirements](#)] embeds an instruction header to the user data packets, and collects the requested data and adds it to the use packet at each network node along the forwarding path. Applications such as path verification, SLA (service-level agreement) assurance can be enabled with iOAM.
- o Control Plane: BGP monitoring protocol (BMP) [[RFC7854](#)] is proposed to monitor BGP sessions and intended to provide a convenient interface for obtaining BGP route views. Data collected using BMP can be further analyzed with big data platforms for network health condition visualization, diagnose and prediction applications.

The general idea of most Telemetry approaches is to collect various information from devices and export to the centralized server for further analysis, and thus providing more network insight. It should not be surprising that any future and even current Telemetry

applications may require the fusion of data acquired from more than one single approach/one single plane. For example, for network troubleshooting purposes, it requires the collection of comprehensive information from devices, such system ID/router ID, interface status, PDUs (protocol data units), device/protocol statistics and so on. Information such as system ID/router ID can be reported by management plane Telemetry approaches, while the protocol related data (especially PDUs) are more fit to be monitored using the control plane Telemetry. With rich information collected in real time at the centralized server, network issues can be localized faster and more accurately, and the root cause analysis can be also provided.

The conventional troubleshooting logic is to log in a faulty router, physically or through Telnet, and by using CLI to display related information/logs for fault source localization and further analysis. There are several concerns with the conventional troubleshooting methods:

1. It requires rich OAM experience for the OAM operator to know what information to check on the device, and the operation is complex;
2. In a multi-vendor network, it requires the understanding and familiarity of vendor specific operations and configurations;
3. Locating the fault source device could be non-trivial work, and is often realized through network-wide device-by-device check, which is both time-consuming and labor-consuming; and finally,
4. The acquisition of troubleshooting data can be difficult under some cases, e.g., when auto recovery is used.

This document proposes the network monitoring for IGP to monitor the running state of IGP, e.g., PDUs, protocol statistics and peer status, which have not been systematically covered by any other Telemetry approach, to facilitate network troubleshooting.

[1.2.](#) Overview

Like BMP, a networking monitoring session is established between each monitored router (NM client) and the NM monitoring station (NM server) through TCP connection. Information are collected directly

from each monitored router and reported to the NM server. The NM message can be both periodic and event-triggered, depending on the message type.

IS-IS [[RFC1195](#)], as one of the most commonly adopted network layer protocols, builds the fundamental network connectivity of an autonomous system (AS). The disfunction of IS-IS, e.g., IS-IS neighbor down, route flapping, MTU mismatch, and so on, could lead to network-wide instability and service interruption. Thus, it is critical to keep track of the health condition of IS-IS, and the availability of information, related to IS-IS running status, is the fundamental requirement. In this document, typical network issues are identified as the use cases of network monitoring. Then the operations and the message formats of network monitoring for IS-IS are defined. Network monitoring for OSPF will be included in the future version.

[2.](#) Terminology

IGP: Interior Gateway Protocol

IS-IS: Intermediate System to Intermediate System

NM: Network Monitoring

IMP: Network Monitoring for IGP

BMP: BGP monitoring protocol

IIH: IS-IS Hello Packet

LSP: Link State Packet

CSNP: Complete Sequence Number Packet

NSNP: Partial Sequence Number Packet

[3.](#) Use Cases

We have identified two typical network issues due to IS-IS disfunction that are currently difficult to detect or localize.

3.1. IS-IS Route Flapping

The IS-IS Route Flapping refers to the situation that one or more routes appear and then disappear in the routing table repeatedly. Route flapping usually comes with massive PDUs interactions (e.g., LSP, LSP purge...), which consume excessive network bandwidth, and excessive CPU processing. In addition, the impact is often network-wide. The localizing of the flapping source and the identifying of root causes haven't been easy work due to various reasons.

The flapping can be caused by system ID conflict, IS-IS neighborship flapping, route source flapping (caused by import route policy misconfiguration) , device clock dis-function with abnormal LSP purge (e.g., 100 times faster) and so on.

- o The system ID conflict check is a network-wide work. If such information is collected centrally to a controller/server, the issues can be identified in seconds, and more importantly, in advance of the actual flapping event.
- o The IS-IS neighborship flapping is typically caused by interface flapping, BFD flapping, CPU high and so on. Conventionally, to located the issue, operators typically identify the target

device(s), and then log in the devices to check related statistics, parsed protocol PDU data and configurations. The manual check often requires a combination of multiple CLIs (check cost/next hop/exit interface/LSP age...) in a repeated manner, which is time-consuming and requires rich OAM experience. If such statistics and configuration data were collected at the server in real-time, the server may analyze them automatically or semi-automatically with troubleshooting algorithms implemented at the server.

- o In the case that route policies are misconfigured, which then causes the route flapping, it's typically difficult to directly identify the responsible policy in a short time. Thus, if the route change history is recorded in correlation with the route policy, then with such record collected at the server, the server

can directly identify the responsible policy with the one-to-one mapping between policy processing and the route attribute change.

- o In the case that flapping comes with abnormal LSP purges, it may be due to continuous LSP corruptions with falsified shorter Remaining Lifetime, or the clock running 100 times faster with 100 times more purge LSPs generated. In order to identify the purge originator, [RFC 6232](#) [RFC6232] proposes to carry the Purge Originator Identification (POI) TLV in IS-IS. However, to analyze the root cause of such abnormal purges, the collection and analysis of LSP PDUs are needed.

[3.2.](#) IS-IS LSDB Synchronization Failure

During the IS-IS flooding, sometimes the LSP synchronization failure happens. The synchronization failure causes can be generally classified into three cases:

- o Case 1, the LSP is not correctly advertised. For example, an LSP sent by Router A fails to be synchronized at Router B. It can be due to incorrect route export policy, or too many prefixes being advertised which exceeds the LSP/MTU threshold, and so on at Router A.
- o Case 2, LSP transmission error, which is typically caused by IS-IS adjacency failure, .e.g., link down/BFD down/authentication failure.
- o Case 3, the LSP is received but not correctly processed. The problem that happens at Router B can be faulty route import policy, or Router B being in Overload mode, or the hardware/software bugs.

With sufficient ISIS PDU related statistics and parsed PDU information recorded at the device, the neighborhood failure in Case 2 can be typically diagnosed at Router A or Router B independently. With such diagnosing information collected (e.g., in the format of reason code) in real-time, the server can identify the root synchronization issue with much less time and labor consumption compared with conventional methods. In Case 1 & 3, the failure is mostly caused by incorrect route policy and software/hardware issue.

By comparing the LSDB with the sent/received LSP, differences can be recognized. Then the difference may further guide the localization of the root cause. Thus, by collecting the LSDBs and sent/received LSPs from the two affected neighbors, the server can have more insights at the synchronization failure.

[4.](#) Message Format

[4.1.](#) Protocol Selection Options

Regarding the network monitoring data export, BMP has been a good option. First of all, BMP serves similar purposes of network monitoring for IGP that reports routes, route statistics and peer status. In addition, BMP has already been implemented in major vendor devices and utilized by operator.

[4.2.](#) Message Types

The variety of IS-IS troubleshooting use cases requires a systematic information report of network monitoring, so that the NM server or any third party analyzer could efficiently utilize the reported messages to localize and recover various network issues. We define NM messages for IS-IS uses the following types:

- o Initiation Message: A message used for the monitored device to inform the NM monitoring station of its capabilities, vendor, software version and so on. For example, the link MTU can be included within the message. The initiation message is sent once the TCP connection between the monitoring station and monitored router is set up. During the monitoring session, any change of the initiation message could trigger an Initiation Message update.
- o Adjacency Status Change Notification Message: A message used to inform the NM monitoring station of the adjacency status change of the monitored device, i.e., from up to down, from down/initiation to up, with possible alarms/logs recorded in the device. This message notifies the NM server of the ongoing IS-IS adjacency change event and possible reasons. If no reason is provided or the provided reason is not specific enough, the NM server can further analyze the IS-IS PDU or the IS-IS statistics.

- o Statistic Report Message: A message used to report the statistics

of the ongoing IS-IS process at the monitored device. For example, abnormal LSP count of the monitored device can be a sign of route flapping. This message can be sent periodically or event triggered. If sent periodically, the frequency can be configured by the operator depending on the monitoring requirement. If it's event triggered, it could be triggered by a counter/timer exceeding the threshold.

- o IS-IS PDU Monitoring Message: A message used to update the NM server of any PDU sent from and received at the monitored device. For example, the IIHs collected from two neighbors can be used for analyzing the adjacency set up failure issue. The LSPs collected from two neighbors can be analyzed for the LSP synchronization issue.
- o Termination Message: A message for the monitored router to inform the monitoring station of why it is closing the NM session. This message is sent when the monitoring session is to be closed.

[4.3.](#) Message Format

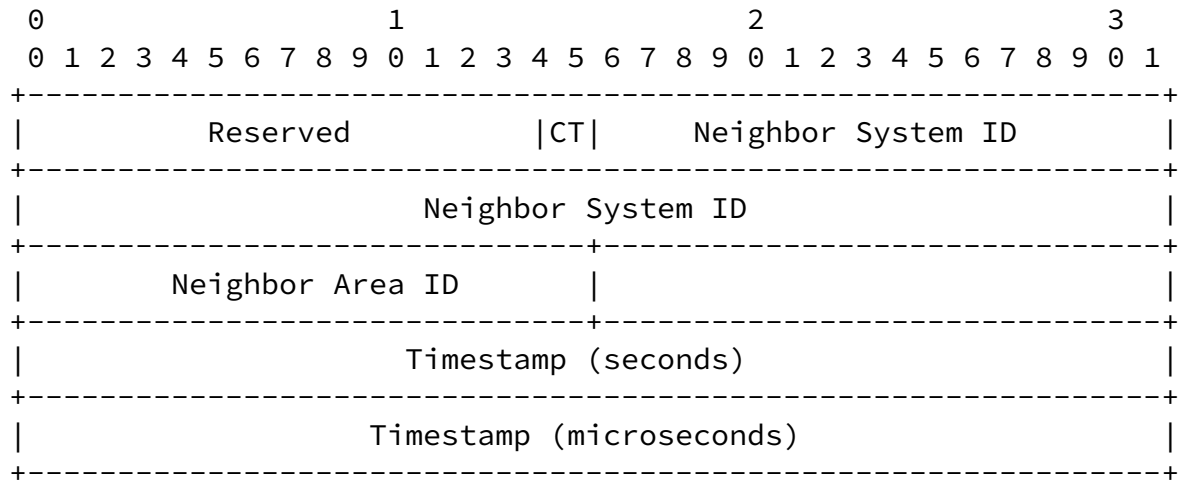
[4.3.1.](#) Common Header

The common header is encapsulated in all messages of network monitoring for IGP. It includes the Version, Message Length and Message Type fields. The common header can reuse the common header of BMP and new message types should be defined for IGP monitoring.

- o Type = TBD: Adjacency Status Change Notification
- o Type = TBD: ISIS Statistic Report
- o Type = TBD: IS-IS PDU Monitoring

[4.3.2.](#) Per Adjacency Header

Except the Initiation and Termination Message, all the rest messages are per adjacency based. Thus, a per adjacency header is defined as follows.



- o Adjacency Flag (2 bytes): The Circuit Type (2 bits) flag specifies if the router is an L1(01), L2(10), or L1/L2(11). If both bits are zeroes (00), the Per Adjacency Header SHALL be ignored. This configuration is used when the statistic is not per-adjacency based, e.g., when reporting the number of adjacencies.
- o Neighbor System ID (6 bytes): identifies the system ID of the remote router.
- o Neighbor Area ID (2 bytes): identifies the area ID of the remote router.
- o Timestamp (4 bytes): records the time when the message is sent/received, expressed in seconds and microseconds since midnight (zero hour), January 1, 1970 (UTC).

[4.3.3.](#) Initiation Message

Three new types of Router Capability TLVs should be defined for IGP monitoring:

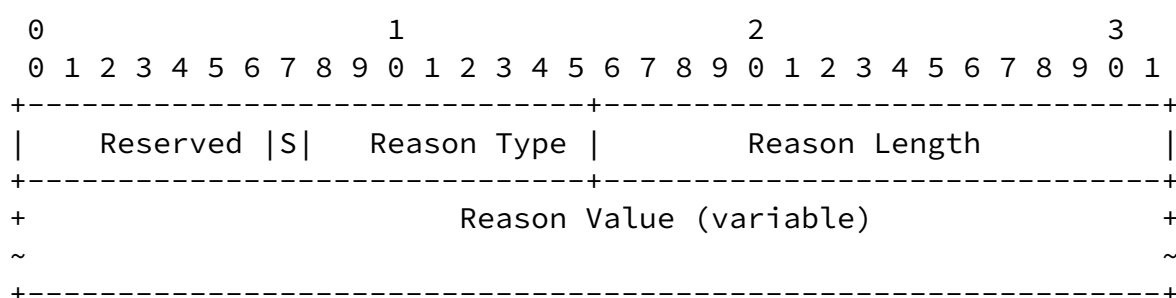
- o Type = TBD: Local System ID. The corresponding Router Capability Value field SHALL indicate the router's System ID
- o Type = TBD: Link MTU. The corresponding Router Capability Value field SHALL indicate the router's link MTU.

[4.3.4.](#) Adjacency Status Change Notification

The Adjacency Status Change Notification Message indicates an IS-IS adjacency status change: from up to down or from initiation/down to up. It consists of the Common Header, Per Adjacency Header and the

Reason TLV. The Notification is triggered whenever the status

changes. The Reason TLV is optional, and is defined as follows. More Reason types can be defined if necessary.



- o Reason Flags (1 byte): The S flag (1 bit) indicates if the Adjacency status is from up to down (set to 0) or from down/initial to up (set to 1). The rest bits of the Flag field are reserved. When the S flag is set to 1, the Reason Type SHALL be set to all zeroes (i.e., Type 0), the Reason Length fields SHALL be set to all zeroes, and the Reason Value field SHALL be set empty.
- o Reason Type (1 byte): indicates the possible reason that caused the adjacency status change. Currently defined types are:
 - * Type = 0: Adjacency Up. This type indicates the establishment of an adjacency. For this reason type, the S flag MUST be set to 1, indicating it's a adjacency-up event. There's no further reason to be provided. The reason Length field SHALL be set to all zeroes, and the Reason Value field SHALL be set empty.
 - * Type = 1: Circuit Down. For this data type, the S flag MUST be set to 0, indicating it's a adjacency-down event. The length field is set to all zeroes, and the value field is set empty.
 - * Type = 2: Memory Low. For this data type, the S flag MUST be set to 0, indicating it's a adjacency-down event. The length field is set to all zeroes, and the value field is set empty.
 - * Type = 3: Hold timer expired. For this data type, the S flag MUST be set to 0, indicating it's a adjacency-down event. The

length field is set to all zeroes, and the value field is set empty.

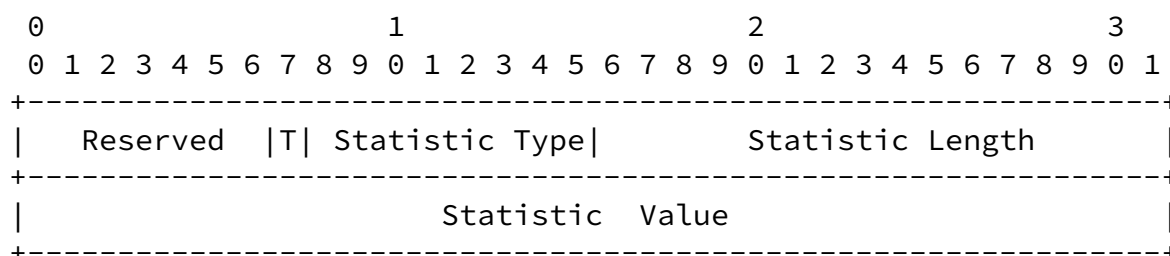
- * Type = 4: String. For this data type, the S flag MUST be set to 0, indicating it's a adjacency-down event. The corresponding Reason Value field indicates the reason specified by the monitored router in a free-form UTF-8 string whose length is given by the Reason Length field.

- o Reason Length (2 bytes): indicates the length of the Reason Value field.
- o Reason Value (variable): includes the possible reason why the Adjacency is down.

[4.3.5.](#) ISIS Statistic Report Message

The ISIS Statistic Report Message reports the statistics of the parameters that are of interest to the operator. The message consists of the Common Header, the Per Adjacency Header and the Statistic TLV. The message include both per-adjacency based statistics and non per- adjacency based statistics. For example, the received/sent LSP counts are per-adjacency based statistics, and the local LSP change times count and the number of established adjacencies are non per- adjacency based statistics. For the non per-adjacency based statistics, the CT Flag (2 bits) in the Per Adjacency Header MUST be set to 00. Upon receiving any message with CT flag set to 00, the Per Adjacency Header SHALL be ignored (the total length of the Per Adjacency Header is 18 bytes as defined in [Section 3.2.2](#), and the message reading/analysis SHALL resume from the Statistic TLV part.

The Statistic TLV is defined as follows.



- o **Statistic Flags (1 byte):** provides information for the reported statistics.
 - * **T flag (1 bit):** indicates if the statistic is for the received-from direction (set to 1) or sent-to direction the neighbor (set to 0)
- o **Statistic Type (1 byte):** specifies the statistic type of the counter. Currently defined types are:
 - * **Type = 0: IIH count.** The T flag indicates if it's a sent or received Hello PDU. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.

- * **Type = 1: Incorrect IIH received count.** For this type, the T flag MUST be set to 1. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.
- * **Type = 2: LSP count.** The T flag indicates if it's a sent or received LSP. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.
- * **Type = 3: Incorrect LSP received count.** For this type, the T flag MUST be set to 1. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.
- * **Type = 4: Retransmitted LSP count.** For this type, the T flag MUST be set to 0. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.
- * **Type = 5: CSNP count.** The T flag indicates if it's a sent or received CSNP. It is a per-adjacency based statistic type, and the CT flag in the Per Adjacency Header MUST NOT be set to 00.
- * **Type = 6: PSNP count.** The T flag indicates if it's a sent or received PSNP. It is a per-adjacency based statistic type, and

the CT flag in the Per Adjacency Header MUST NOT be set to 00.

- * Type = 7: Number of established adjacencies. It's a non per-adjacency based statistic type, and thus for the monitoring station to recognize this type, the CT flag in the Per Adjacency Header MUST be set to 00.
- * Type = 8: LSP change time count. It's a non per-adjacency based statistic type, and thus for the monitoring station to recognize this type, the CT flag in the Per Adjacency Header MUST be set to 00.
- o Statistic Length (2 bytes): indicates the length of the Statistic Value field.
- o Statistic Value (4 bytes): specifies the counter value, which is a non-negative integer.

[4.3.6.](#) IS-IS PDU Monitoring Message

The IS-IS PDU Monitoring Message is used to update the monitoring station of any PDU sent from and received at the monitored device per neighbor. Following the Common Header and the Per Adjacency Header

is the IS-IS PDU. To tell whether it's a sent or received PDU, the monitoring station can analyze the source and destination addresses in the reported PDUs.

[4.3.7.](#) Termination Message

This document does not change the Termination Message defined by [RFC7854](#).

[5.](#) IANA

TBD

[6.](#) Contributors

TBD

[7.](#) Acknowledgments

TBD

8. 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., Lapukhov, P., and r. remy@barefootnetworks.com, "Requirements for In-situ OAM", [draft-brockners-inband-oam-requirements-03](#) (work in progress), March 2017.

[I-D.chen-npm-use-cases]

Chen, H., Li, Z., Xu, F., Gu, Y., and Z. Li, "Network-wide Protocol Monitoring (NPM): Use Cases", [draft-chen-npm-use-cases-00](#) (work in progress), March 2019.

[I-D.ietf-netconf-yang-push]

Clemm, A. and E. Voit, "Subscription to YANG Datastores", [draft-ietf-netconf-yang-push-25](#) (work in progress), May 2019.

[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.

Gu, et al.

Expires August 23, 2021

[Page 13]

Internet-Draft

Network Monitoring For IGP

February 2021

[I-D.song-ntf]

Song, H., Zhou, T., Li, Z., Fioccola, G., Li, Z., Martinez-Julia, P., Ciavaglia, L., and A. Wang, "Toward a Network Telemetry Framework", [draft-song-ntf-02](#) (work in progress), July 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>>.

- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", [RFC 1195](#), DOI 10.17487/RFC1195, December 1990, <<https://www.rfc-editor.org/info/rfc1195>>.
- [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>>.
- [RFC6232] Wei, F., Qin, Y., Li, Z., Li, T., and J. Dong, "Purge Originator Identification TLV for IS-IS", [RFC 6232](#), DOI 10.17487/RFC6232, May 2011, <<https://www.rfc-editor.org/info/rfc6232>>.
- [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>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", [RFC 7752](#), DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [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>>.

Authors' Addresses

Gu, et al.

Expires August 23, 2021

[Page 14]

Internet-Draft

Network Monitoring For IGP

February 2021

Yunan Gu
Huawei
156 Beiqing Road
Beijing, 100095

China

Email: guyunan@huawei.com

Shuanglong Chen
Huawei
156 Beiqing Road
Beijing, 100095
China

Email: chenshuanglong@huawei.com

Yingzhen Qu
Futurewei
United States

Email: yingzhen.qu@futurewei.com

Huanan Chen
China Telecom
109 West Zhongshan Ave
Guangzhou
China

Email: chenhuanan@189.com

Zhenbin Li
Huawei
156 Beiqing Road
Beijing, 100095
China

Email: lizhenbin@huawei.com