OPSAWG
Internet-Draft
Intended status: Informational
Expires: March 9, 2020

H. Song, Ed.
Futurewei
Z. Li
T. Zhou
Huawei
F. Qin
China Mobile
J. Shin
SK Telecom
J. Jin
LG U+

September 6, 2019

In-situ Flow Information Telemetry Framework
draft-song-opsawg-ifit-framework-04

Abstract

Unlike the existing active and passive OAM techniques, the emerging on-path flow telemetry techniques provide unmatched visibility into user traffic, showing great application potential not only for today's network OAM but also for future's automatic network operation. Summarizing the current industry practices that addresses the deployment challenges and application requirements, we provide a closed-loop framework, named In-situ Flow Information Telemetry (iFIT), for efficiently applying a family of underlying on-path flow telemetry techniques in various network environments. The framework enumerates several key architectural components and describes how these components are assembled together to achieve a complete and closed-loop working solution for on-path flow telemetry. Following such a framework allows better scalability, fosters application innovations, and promotes both vertical and horizontal interoperability.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

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

This Internet-Draft will expire on March 9, 2020.

Copyright Notice

Table of Contents

1. Requirements and Challenges

Application-aware network operation is important for user SLA compliance, service path enforcement, fault diagnosis, and network resource optimization. A family of on-path flow telemetry techniques, including In-situ OAM (IOAM) [I-D.brockners-inband-oam-data], PBT [I-D.song-ippm-postcard-based-telemetry], IFA [I-D.kumar-ippm-ifa], Enhanced AM [I-D.zhou-ippm-enhanced-alternate-marking], and HTS [I-D.mirsky-ippm-hybrid-two-step], are emerging, which can provide flow information on the entire forwarding path on a per-packet basis in real time. These techniques are very different from the previous active and passive OAM schemes in that they directly modify the user packets and can gain visibility on every user packet. Given the unique characteristics of such techniques, we categorize these on-path telemetry techniques as the hybrid OAM type III, supplementing the classification defined in [RFC7799].

These techniques are invaluable for application-aware network operations not only in data center and enterprise networks but also in carrier networks which may cross multiple domains. Carrier network operators have shown strong interests in utilizing such techniques for various purposes. For example, it is vital for the operators who offer the bandwidth intensive, latency and loss sensitive services such as video streaming and gaming to closely monitor the relevant flows in real time as the indispensable first step for any further measure.

However, successfully applying such techniques in carrier networks poses several practical challenges:

- C1: On-path flow telemetry incurs extra packet processing which may strain the network data plane. The potential impact on the forwarding performance creates an unfavorable "observer effect" which not only damages the fidelity of the measurement but also defies the purpose of the measurement.

- C2: On-path flow telemetry can generate a huge amount of OAM data which may claim too much transport bandwidth and inundate the servers for data collection, storage, and analysis. Increasing the data handling capacity is technically viable but expensive.

- C3: The collectible data defined currently are essential but limited. As the network operation evolves to become intent-based and automatic, and the trends of network virtualization, network convergence, and packet-optical integration continue, more data will be needed in an on-demand and interactive fashion.

Flexibility and extensibility on data defining and acquiring must be considered.

- C4: If we were to apply some on-path telemetry technique in today's carrier networks, we must provide solutions to tailor the provider's network deployment base and support an incremental deployment strategy. That is, we need to come up with encapsulation schemes for various predominant protocols such as Ethernet, IPv4, and MPLS with backward compatibility and properly handle various transport tunnels.

- C5: Applying only a single underlying telemetry technique may lead to defective result. For example, packet drop can cause the lost of the flow telemetry data and the packet drop location and reason remains unknown if only In-situ OAM trace option is used. A comprehensive solution needs the flexibility to switch between different underlying techniques and adjust the configurations and parameters at runtime.

2. iFIT Framework Overview

To address these challenges, we propose a framework based on multiple network operators' requirements and the common industry practice, which can help to build a workable on-path flow telemetry solution. We name the framework "In-situ Flow Information Telemetry" (iFIT) to reflect the fact that this framework is dedicated to the on-path telemetry data about user/application flow experience. As a solution framework, iFIT works a level higher than any specific OAM techniques, be it active, passive, or hybrid. The framework is built up on a few architectural components. By assembling these components together, a closed-loop is formed to provide a complete solution for a particular static, dynamic, and interactive telemetry applications.

iFIT is an open framework. It does not enforce any implementation details for each component. Users are free to pick one or more underlying techniques and design their own algorithms and architectures to fit in each component and make all the components work in concert.

The network architecture that applies iFIT is shown in Figure 1. The iFIT domain is confined between the iFIT head nodes and the iFIT end nodes. An iFIT domain may cross multiple network domains. iFIT support two basic on-path telemetry data collection modes: passport mode (e.g., IOAM trace option and IFA), in which telemetry data are carried in user packets and exported at the iFIT end nodes, and postcard mode (e.g., PBT), in which each node in the iFIT domain may export telemetry data through independent OAM packets. Note that the

boundary between the two modes can be blurry. An application only need to mix the two modes.

The key components of iFIT is listed as follows:

- Smart flow and data selection policy to address C1.
- Export data reduction to address C2.
- Dynamic network probe to address C3.
- Encapsulation and tunnel modes to address C4.
- On-demand technique selection to address C5.

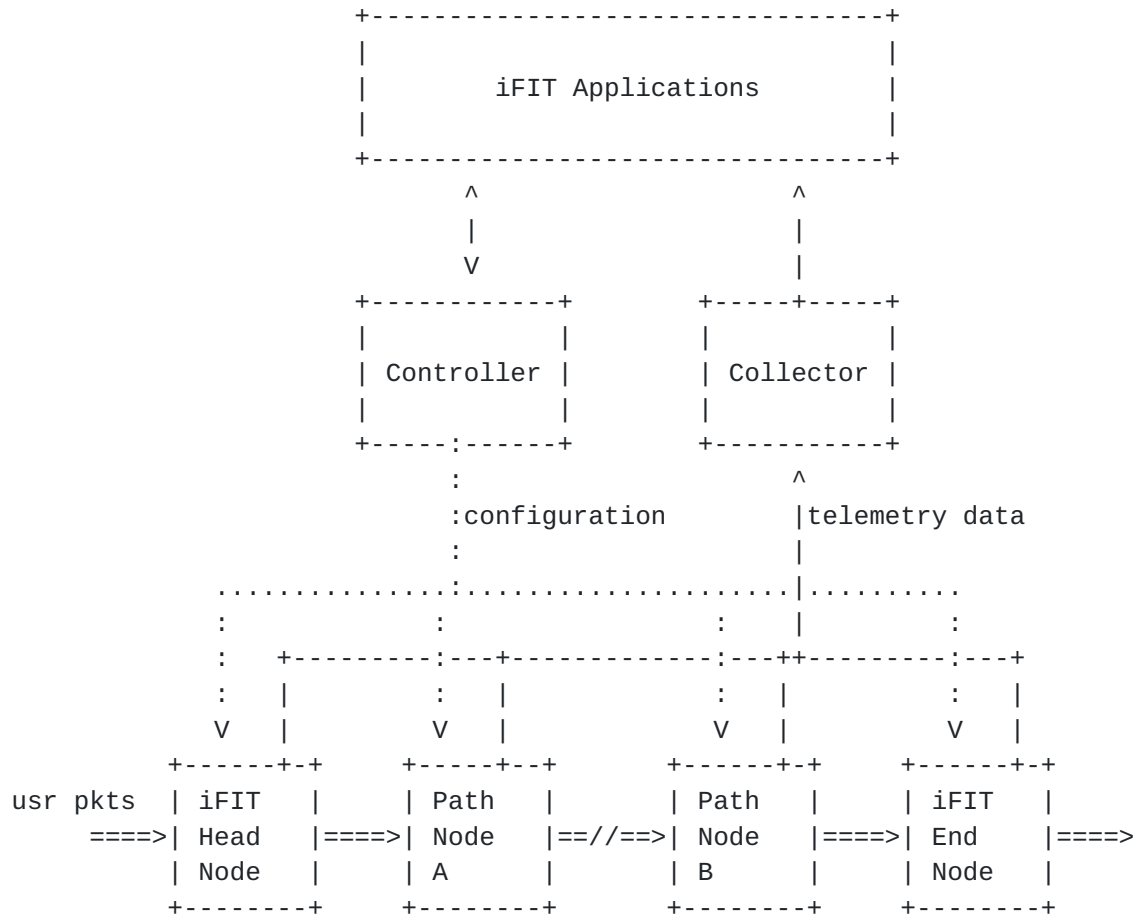

Figure 1: iFIT Architecture

In the remaining of the document, we provide the detailed discussion of the iFIT's components.

3. Smart Flow and Data Selection

In most cases, it is impractical to enable the data collection for all the flows and for all the packets in a flow due to the potential performance and bandwidth impacts. Therefore, a workable solution must select only a subset of flows and flow packets to enable the data collection, even though this means the loss of some information.

In data plane, the Access Control List (ACL) provides an ideal means to determine the subset of flow(s). [I-D.song-ippm-ioam-data-validation-option] describes how one can set a sample rate or probability to a flow to allow only a subset of flow packets to be monitored, how one can collect different set of data for different packets, and how one can disable or enable data collection on any specific network node. The document further introduces enhancement to IOAM to allow any node to accept or deny the data collection in full or partially.

Based on these flexible mechanisms, iFIT allows applications to apply smart flow and data selection policies to suit the requirements. The applications can dynamically change the policies at any time based on the network load, processing capability, focus of interest, and any other criteria. We have developed some adaptive algorithm which can limit the performance impact and yet achieve the satisfactory telemetry data density.

4. Export Data Reduction

The flow telemetry data can catch the dynamics of the network and the interactions between user traffic and network. Nevertheless, the data inevitably contain redundancy. It is advisable to remove the redundancy from the data in order to reduce the data transport bandwidth and server processing load.

In addition to efficiently encode the export data (e.g., IPFIX [RFC7011] or protobuf [1]), iFIT can also cache the data and send the accumulated data in batch if the data is not time sensitive. Various deduplication and compression techniques can be applied on the batch data.

From the application perspective, an application may only be interested in some special events which can be derived from the telemetry data. For example, in case that the forwarding delay of a packet exceeds a threshold or a flow changes its forwarding path is of interest, it is unnecessary to send the original raw data to the

data collecting and processing servers. Rather, iFIT takes advantage of the in-network computing capability of network devices to process the raw data and only push the event notifications to the subscribing applications.

5. Dynamic Network Probe

Due to the limited data plane resource, it is unlikely one can provide all the data all the time. On the other hand, the data needed by applications may be arbitrary but ephemeral. It is critical to meet the dynamic data requirements with limited resource.

Fortunately, data plane programmability allows iFit to dynamically load new data probes. These on-demand probes are called Dynamic Network Probes (DNP) [I-D.song-opsawg-dnp4iq]. DNP is the technique to enable probes for customized data collection in different network planes. When working with IOAM or PBT, DNP is loaded to the data plane through incremental programming or configuration. The DNP can effectively conduct data generation, processing, and aggregation.

DNP introduces enough flexibility and extensibility to iFIT. It can implement the optimizations for export data reduction motioned in the previous section. It can also generate custom data as required by today and tomorrow's applications.

6. Encapsulation and Tunnel Modes

Since MPLS and IPv4 network are still prevalent in carrier networks. iFIT provides solutions to apply the on-path flow telemetry techniques in such networks. PBT-M [I-D.song-ippm-postcard-based-telemetry] does not introduce new headers to the packets so the trouble of encapsulation for a new header is avoided. In case a technique that requires a new header is preferred, [I-D.song-mpls-extension-header] provides a means to encapsulate the extra header using an MPLS extension header. As for IPv4, it is possible to encapsulate the new header in an IP option. For example, RAO [RFC2113] can be used to indicate the presence of the new header. A recent proposal [I-D.herbert-ipv4-eh] that introduces the IPv4 extension header may lead to a long term solution.

In carrier networks, it is common for user traffic to traverse various tunnels for QoS, traffic engineering, or security. iFIT supports both the uniform mode and the pipe mode for tunnel support as described in [I-D.song-ippm-ioam-tunnel-mode]. With such flexibility, the operator can either gain a true end-to-end visibility or apply a hierarchical approach which isolates the monitoring domain between customer and provider.

7. On-demand Technique Selection and Integration

With multiple underlying data collection and export techniques at its disposal, iFIT can flexibly adapt to different network conditions and different application requirements.

For example, depending on the types of data that are of interest, iFIT may choose either IOAM or PBT to collect the data; if an application needs to track down where the packets are lost, it may switch from IOAM to PBT.

iFIT can further integrate multiple data plane monitoring and measurement techniques together and present a comprehensive data plane telemetry solution to network operating applications.

8. Summary and Future Work

iFIT is a framework for applying on-path data plane telemetry techniques. Combining with algorithmic and architectural schemes that fit into the framework components, iFIT framework enables a practical telemetry solution based on two basic on-path traffic data collection modes: passport and postcard.

The operation of iFIT differs from both active OAM and passive OAM as defined in [RFC7799]. It does not generate any active probe packets or passively observe unmodified user packets. Instead, it modifies selected user packets to collect useful information about them. Therefore, the iFIT operation can be considered the hybrid type III mode, which can provide more flexible and accurate network OAM.

More challenges and corresponding solutions for iFIT may need to be covered. For example, how iFIT can fit in the big picture of autonomous networking and support closed control loops. A complete iFIT framework should also consider the cross-domain operations. We leave these topics for future revisions.

9. Security Considerations

No specific security issues are identified other than those have been discussed in the drafts on on-path flow information telemetry.

10. IANA Considerations

This document includes no request to IANA.

11. Contributors

TBD.

12. Acknowledgments

TBD.

13. References

13.1. 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>.

[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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

13.2. Informative References

[I-D.brockners-inband-oam-data] Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., and d. daniel.bernier@bell.ca, "Data Fields for In-situ OAM", draft-brockners-inband-oam-data-07 (work in progress), July 2017.

[I-D.herbert-ipv4-eh] Herbert, T., "IPv4 Extension Headers and Flow Label", draft-herbert-ipv4-eh-01 (work in progress), May 2019.

[I-D.kumar-ippm-ifa] Kumar, J., Anubolu, S., Lemon, J., Manur, R., Holbrook, H., Ghanwani, A., Cai, D., Ou, H., and L. Yizhou, "Inband Flow Analyzer", draft-kumar-ippm-ifa-01 (work in progress), February 2019.

[I-D.mirsky-ippm-hybrid-two-step] Mirsky, G., Lingqiang, W., and G. Zhui, "Hybrid Two-Step Performance Measurement Method", draft-mirsky-ippm-hybrid-two-step-03 (work in progress), April 2019.

[I-D.song-ippm-ioam-data-validation-option]
Song, H. and T. Zhou, "In-situ OAM Data Validation Option", draft-song-ippm-ioam-data-validation-option-02 (work in progress), April 2018.

[I-D.song-ippm-ioam-tunnel-mode]
Song, H., Li, Z., Zhou, T., and Z. Wang, "In-situ OAM Processing in Tunnels", draft-song-ippm-ioam-tunnel-mode-00 (work in progress), June 2018.

[I-D.song-ippm-postcard-based-telemetry]
Song, H., Zhou, T., Li, Z., Shin, J., and K. Lee, "Postcard-based On-Path Flow Data Telemetry", draft-song-ippm-postcard-based-telemetry-04 (work in progress), June 2019.

[I-D.song-mpls-extension-header]
Song, H., Li, Z., Zhou, T., and L. Andersson, "MPLS Extension Header", draft-song-mpls-extension-header-02 (work in progress), February 2019.

[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-ippm-enhanced-alternate-marking]
Zhou, T., Fioccola, G., Li, Z., Lee, S., Cociglio, M., and Z. Li, "Enhanced Alternate Marking Method", draft-zhou-ippm-enhanced-alternate-marking-03 (work in progress), July 2019.

[RFC2113] Katz, D., "IP Router Alert Option", RFC 2113, DOI 10.17487/RFC2113, February 1997, <https://www.rfc-editor.org/info/rfc2113>.

[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>.

[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>.

13.3. URIs

[1] https://developers.google.com/protocol-buffers/

Authors' Addresses

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

Email: haoyu.song@futurewei.com

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

Email: lizhenbin@huawei.com

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

Email: zhoutianran@huawei.com

Fengwei Qin
China Mobile
No. 32 Xuanwumenxi Ave., Xicheng District
Beijing, 100032
P.R. China

Email: qinfengwei@chinamobile.com

Jongyoon Shin
SK Telecom
South Korea

Email: jongyoon.shin@sk.com

Jaewhan Jin
LG U+
South Korea

Email: daenamu1@lguplus.co.kr