ETSI GR ENI 012 V1.1.1 (2022-03)

Experiential Networked Intelligence (ENI); Reactive In-situ Flow Information Telemetry

Disclaimer

The present document has been produced and approved by the Experiential Networked Intelligence (ENI) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGR/ENI-0022_Flow_Info_Tele

Keywords
artificial intelligence, network, telemetry

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) Experiential Networked Intelligence (ENI).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document describes the motivation, issues, and challenges of using flow-oriented on-path telemetry techniques which provide relevant measurement or event reports to the AI-enabled network entities.

The present document outlines a reference framework, named as "In-situ Flow Information Telemetry (IFIT)" and identifies technical issues, including modes of flow-oriented on-path telemetry; IFIT-based reactive telemetry framework and technical issues, including intelligent flow and packet selection, intelligent data export, dynamic network probe, on-demand underlying technique selection.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] IETF RFC 7276 (June 2014): "An Overview of Operations, Administration, and Maintenance (OAM) Tools".

[i.2] IETF RFC 4443 (March 2006): "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification".

[i.3] IETF RFC 5880 (June 2010): "Bidirectional Forwarding Detection (BFD)".

[i.4] IETF RFC 7799 (May 2016): "Active and Passive Metrics and Methods (with Hybrid Types In-Between)".

[i.5] IETF RFC 4656 (September 2006): "A One-way Active Measurement Protocol (OWAMP)".

[i.6] IETF RFC 5357 (October 2008): "A Two-Way Active Measurement Protocol (TWAMP)".

[i.7] IETF RFC 7011 (September 2013): "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information".

[i.8] IETF RFC 8321 (January 2018): "Alternate-Marking Method for Passive and Hybrid Performance Monitoring".

[i.9] draft-ietf-ippm-ioam-data (June 2021): "Data Fields for In-situ OAM".

[i.10] IETF draft-ietf-ippm-ioam-direct-export (August 2021): "In-situ OAM Direct Exporting".

[i.11] IETF RFC 8889 (August 2020): "Multipoint Alternate-Marking Method for Passive and Hybrid Performance Monitoring".

[i.12] IETF draft-song-opsawg-ifit-framework (work in progress): "In-situ Flow Information Telemetry".

[i.13] Bo Lu, Ling Xu, Yuezhong Song, Longfei Dai, Min Liu, Tianran Zhou, Zhenbin Li and Haoyu Song: "IFIT: Intelligent Flow Information Telemetry". In Proceedings of the ACM SIGCOMM 2019 Conference Posters and Demos (SIGCOMM Posters and Demos '19). Association for Computing Machinery, New York, NY, USA, p15-17.

NOTE: Available at https://doi.org/10.1145/3342280.3342292.

[i.14] ETSI GR ENI 009 (V1.1.1): "Experiential Networked Intelligence (ENI); Definition of Data Processing Mechanisms".

[i.15] IETF RFC 8639 (September 2019): "Subscription to YANG Notifications".

[i.16] IETF RFC 8640 (September 2019): "Dynamic Subscription to YANG Events and Datastores over NETCONF".

[i.17] IETF RFC 8641 (September 2019): "Subscription to YANG Notifications for Datastore Updates".

[i.18] IETF RFC 8650 (November 2019): "Dynamic Subscription to YANG Events and Datastores over RESTCONF".

[i.19] draft-ietf-ippm-ioam-yang (work in progress): "A YANG Data Model for In-Situ OAM".

[i.20] IETF RFC 7950 (August 2016): "The YANG 1.1 Data Modeling Language".

[i.21] IETF RFC 6241 (June 2011): "Network Configuration Protocol (NETCONF)".

[i.22] IETF RFC 8040 (January 2017): "RESTCONF Protocol".

[i.23] draft-ietf-idr-sr-policy-ifit: "BGP SR Policy Extensions to Enable IFIT".

[i.24] draft-chen-pce-pcep-ifit (work in progress): "Path Computation Element Communication Protocol (PCEP) Extensions to Enable IFIT".

[i.25] ETSI GS ENI 005 (V2.1.1): "Experiential Networked Intelligence (ENI); System Architecture".

[i.26] IETF RFC 5475 (March 2009): "Sampling and Filtering Techniques for IP Packet Selection".

[i.27] Shuping Peng, Jianwei Mao, Ruizhao Hu and Zhenbin Li: "Demo Abstract: APN6: Application-aware IPv6 Networking", IEEE INFOCOM 2020.

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

flow-oriented on-path telemetry: specific class of network forwarding-plane telemetry techniques, including In-situ OAM (IOAM), Enhanced Alternate Marking (EAM), Postcard-Based Telemetry (PBT), and Hybrid Two Steps (HTS)

In-situ Flow Information Telemetry（IFIT): network OAM data plane on-path telemetry techniques, including In-situ OAM (IOAM), Direct Exporting (DEX IOAM (IOAM-DEX), Postcard-Based Telemetry (PBT), and Alternate Marking

NOTE 1: It can provide flow information on the entire forwarding path on a per-packet basis in real time. "In-situ" is Latin which can be translated as "in the original place".

NOTE 2: See IETF RFC 8321 [i.8].

reactive telemetry: telemetry operation in a dynamic and interactive fashion

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ACL	Access Control List
AMF	Access and Mobility Management Function
AMM	Enhanced Alternate Marking Method
API	Application Programming Interface
APN	APplication-aware Network
ASG	Aggregation Site Gateway
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
BUM	Broadcast, Unknown-Unicast and Multicast
DEX	Direct Exporting
DNP	Dynamic Network Probes
E2E	End-to-End
EAM	Enhanced Alternate Marking
ECMP	Equal-Cost Multipath
ENI	Experiential Networked Intelligence
GPB	Google Protocol Buffer
GPRS	General Packet Radio Service
GTP	GPRS Tunnelling Protocol
HD	High-Definition
HTS	Hybrid Two Steps
IFIT	In-situ Flow Information Telemetry
IOAM	In-situ OAM
IP	Internet Protocol
IPFIX	IP Flow Information eXport
IPFPM	IP Flow Performance Measurement
IPPM	IP Performance Measurement
MDT	Model Driven Telemetry
NBI	North Bound Interface
NMS	Network Management System
NP	Network Processor
OAM	Operation, Administration and Maintenance
OTT	Over The Top
OWAMP	One-Way Active Measurement Protocol
PBT	Postcard-Based Telemetry
PCEP	Path Computation Element communication Protocol
PM	Performance Management
RSG	Radio Service Gateway
SBI	South Bound Interface
SCTP	Stream Control Transmission Protocol
SDN	Software-Defined Network
SD-WAN	Software-defined WAN
SLA	Service Level Agreement
SR	Segment Routing
TM	Traffic Manager
TWAMP	Two-Way Active Measurement Protocol
UPF	User Plane Function
VPN	Virtual Private Network
WAN	Wide Area Network
WG	Working Group
YANG	Yet Another Next Generation

4 Introduction

4.1 Background, Motivation and Benificial Aspests

4.1.1 Background

This clause presents background and deployment challenges of flow-oriented on-path telemetry, as well as the components of a reference framework.

As introduced in ETSI GS ENI 005 [i.25], current network management and performance measurement functions are not optimized due to the different technologies and implementations from different vendors. The human-machine interaction challenges increase the time to market of innovative and advanced services (including the new performance management tools).

In-situ Flow Information Telemetry (IFIT) is a family of hybrid data-plane telemetry technologies, through which the flow quality measurement information is directly recorded and encapsulated in data packets to implement flow quality visualization at a granularity of each data packet. This flow quality measurement information that may be carried on a complete forwarding path in real time at a per-packet granularity may include device and interface information as well as a nanosecond-precision cache time of a packet in each network device as well as identification contention queue flow information. In-situ flow information telemetry technologies include e.g. In-situ OAM (IOAM) [i.9], IOAM Direct Exporting (IOAM-DEX) [i.10], and Alternate Marking Method (AMM) [i.8]. This family of In-situ flow information telemetry technologies are currently defined by IETF.

4.1.2 Motivation

Currently there is no efficient and extensible standard-based mechanism to provide smart, context-aware and flexible performance management. In addition, Performance Measurement tools should adapt to variations in network conditions, changes in user needs and business goals. All this will be possible by using In-situ flow information telemetry techniques. Moreover, this approach further enables the use of real-time closed control loops and also helps to optimize network resources through the use of automation and smart application of network monitoring. Network intelligence allows to start without examining in depth and, if there are problems in some network portions these are detected. Then, it can be possible to determine which parts of the network are affected and start an in-depth analysis only where and when is necessary. The resulting telemetry information can be used to understand what is going on and to eventually try to solve it in order to maintain Service Level Agreements (SLAs).

ETSI GS ENI 005 [i.25] defines a Functional Block architecture that helps to address the application of In-situ flow information telemetry. The experiential architecture and self-learning principle are key to implement a smart, context-aware and flexible performance management.

4.1.3 Benificial Aspests

Efficient network OAM increasingly depends on high-quality visualization of network data plane quality. Traditional OAM technologies are widely used, including network fault detection, network fault isolation, network fault reporting, and network performance measurement, IETF RFC 7276 [i.1]. For example, traditional IP Ping, IETF RFC 4443 [i.2] and Bidirectional Forwarding Detection (BFD), IETF RFC 5880 [i.3] are used for connectivity detection. According to IETF RFC 7799 [i.4], performance measurement can be classified into three types, aka active performance measurement, passive performance measurement, and hybrid performance measurement. For example, Two-Way Active Measurement Protocol (TWAMP), IETF RFC 5357 [i.6], is a typical active performance measurement method that operates by injecting proactive probe packets to measure the loss and delay.

Different from active performance measurement, passive performance measurement directly monitors data flows without sending additional probe packets or modifying data packets. For example, the IP Flow Information eXport (IPFIX) protocol [i.7] may transmit IP data flow statistics from a device to a collector by using a pre-defined data output format. In addition, hybrid performance measurement combines active performance measurement and passive performance measurement to modify certain fields of data packets without introducing additional probe packets to the network. For example, IP Flow Performance Measurement (IPFPM) [i.8] directly monitors real data flows by colouring data packets, which is a typical hybrid performance measurement technology. Because the hybrid performance measurement method does not introduce additional probe packets, the accuracy of the performance measurement can be equal to that of the passive measurement.

Traditional network performance measurement technologies (such as TWAMP [i.6] and IPFIX [i.7]) cannot meet the requirements of high-precision and real-time network performance monitoring. A new measurement technology is required to meet the requirements of future network and service development. In addition, intelligence has become the developing trend of network. The new APplication-aware Network (APN) [i.27] architecture describes the ability of the network to acquire and manage current information about users and applications. This information can be used to optimize the use of network resources and improve the quality of service. In addition, the emerging In-situ Flow Information Telemetry (IFIT) technologies provide high-precision visualization of flow quality (such as jitter, delay, packet loss).

Although In-situ flow information telemetry is beneficial, it is to be addressed in the following practical deployment challenges. First, In-situ flow information telemetry incurs extra packet processing which may cause stress on the network data plane. The potential impact on the forwarding performance creates an unfavourable "observer effect". For example, the growing IOAM data per hop can negatively affect service levels by increasing the serialization delay and header parsing delay. Second, In-situ flow information telemetry can generate a considerable amount of data which may consume too much transport bandwidth and the servers used for data collection, storage, and analysis, may collapse. For example, if IOAM is applied to all the traffic, one node may collect a few tens of bytes as telemetry data per packet. Third, as the network operation evolves to be fully automated, and the trends of network virtualization and packet-optical integration continue, more data is needed in an on-demand and interactive fashion. Flexibility and extensibility on data defining, aggregation, acquisition, and filtering, is to be considered. Lastly, as applying only a single underlying in-situ flow information telemetry technique may lead to a defective result, for example, packet drop can cause the loss of the flow telemetry data. Therefore, the reason why the packet drop has occurred remains unknown if only the IOAM trace option [i.9] is used. As such, a comprehensive solution needs the flexibility to switch between different underlying techniques and adjust the configurations and parameters at runtime. Hence, system-level management is needed.

The present document provides an IFIT-based reactive Telemetry framework, which addresses the aforementioned handicaps for deployment. By following this framework, an effectively and implementable automatic data flow quality measurement scheme for data flow becomes possible. IFIT-based reactive Telemetry framework requires intelligent flow selection, efficient data handling, dynamic network probe, and tunnel encapsulation to enable on-demand network performance measurement.

4.2 Modes of Flow-oriented On-path Telemetry

This clause lists various telemetry techniques in the category of flow-oriented on-path telemetry, such as In-situ OAM (IOAM) [i.9], IOAM Direct Exporting (IOAM-DEX) [i.10], Alternate Marking Method (AMM) [i.8], and classifies various modes of flow-oriented on-path telemetry in accordance to IETF RFC 7799 [i.4].

Traditional OAM technologies are widely used in network OAM and management, including network fault detection, network fault isolation, network fault reporting, and network performance detection. According to IETF RFC 7799 [i.4], performance measurement can be classified into three types: active performance measurement, passive performance measurement, and hybrid performance measurement.

In proactive performance measurement, probe packets are sent on the network, and the network performance is deduced by measuring the probe packets. Active performance measurement methods, such as IP Ping [i.2], Bidirectional Forwarding Detection (BFD) [i.3], One-Way Active Measurement Protocol (OWAMP) [i.5], and Two-Way Active Measurement Protocol (TWAMP) [i.6].

Different from active performance measurement, passive performance measurement obtains performance parameters by directly monitoring service data flows. No additional detection packets are sent or service packets need to be modified. Therefore, the performance can be accurately and accurately reflected. Passive performance measurement methods, such asIP Flow Information Export (IPFIX) [i.7], which is a statistical and output standard for IP data flows. IP data flow statistics can be transmitted from one output to the collector through a defined data output format.

Hybrid performance measurement is a combination of active performance measurement and passive performance measurement. It does not need to send additional probe packets on the network. Instead, it only needs to modify some fields of service packets to measure network performance. Because the hybrid performance measurement method does not introduce additional active measurement packets, the accuracy of the performance measurement can be equal to that of the passive measurement, hybrid performance measurement methods, such as IOAM [i.9], IOAM-DEX [i.10], Alternate Marking [i.8], which directly monitors real data flows by inserting instruction header or colouring data packets.

5 Overview of IFIT

5.1 IFIT-based Reactive Telemetry Framework

As a hybrid performance measurement technology, the IFIT technology provides high-precision visualization of flow quality and real-time network fault alarms (such as jitter, delay, packet loss) to meet the requirements for high-performance network quality measurement in the future. IFIT encapsulates flow quality measurement information into user data packets to implement real-time and per-packet flow quality measurement.

Figure 1 shows an IFIT-based reactive telemetry framework, which includes Application and Management System, Controller, and IFIT-enabled forwarding devices.

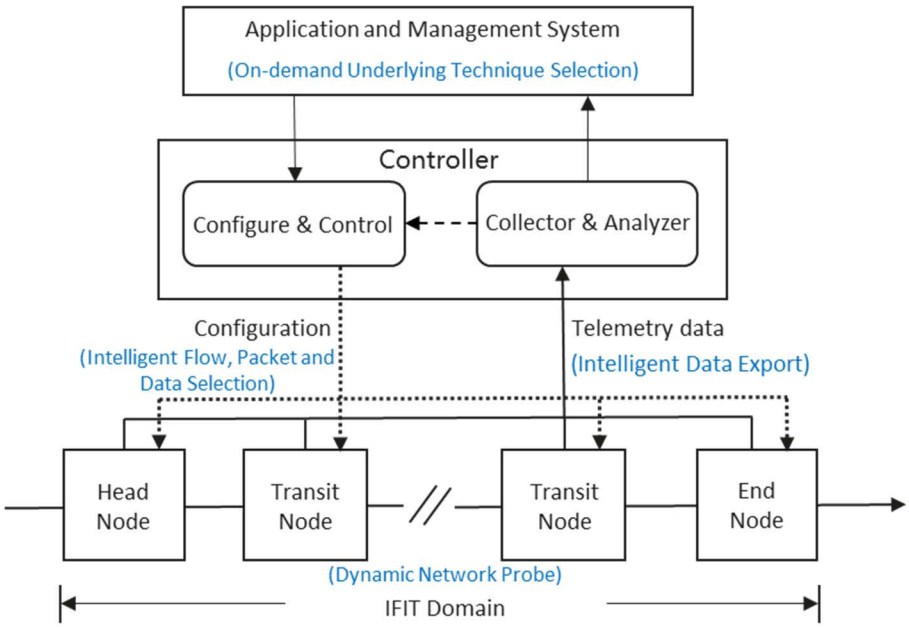

Figure 1: IFIT-based Reactive Telemetry Framework

As shown in the IFIT-based reactive telemetry framework, to meet the measurement requirements of different applications, multiple data-plane measurement technologies and data exporting technologies can be flexibility integrated to provide comprehensive performance information for network OAM. For example, for different types of information data, IOAM or Alternate Marking may be selected to collect information. In addition, switching from the IOAM mode [i.9] to the IOAM-DEX mode [i.10] for fault location. After the telemetry data is processed and analysed, the analysing results may be used to instruct the controller to modify a configuration of a node in the IFIT domain for adjusting data collected by the IFIT. Therefore, the process may be dynamic and interactive.

The IFIT domain can cross multiple network domains. The nodes that enter and leave the IFIT domain are called the Head Node and End Node. The ingress node is responsible for encapsulating the IFIT instruction header into data packets. All nodes in the IFIT domain can perform the specified IFIT function. The end node is to be able to capture all packets with IFIT headers and metadata, remove the IFIT headers and IFIT metadata to ensure that any data packet with IFIT-specific headers and metadata does not leak out of the IFIT domain, and then forward them out of the IFIT field.

In the IFIT-based Reactive Telemetry Framework shown in Figure 1, each functional components are as follows:

a) The Application and Management System is responsible for inputting OAM measurement intent and displaying measurement analysis results. On the one hand, the intent of network quality measurement from service applications and OAM systems is received, converted into network configuration policies, and delivered to the controller. The IFIT network configuration policy generated by the application and management system, which includes information such as a specified flow object to be measured, a performance indicator to be collected, and a test data exporting mode (passport mode or postcard mode). On the other hand, the application and management system receives IFIT quality measurement data and analysis results from the collector and analyser, then displays the results in a visualized manner.

b) The Controller consists of two functional components: Configuration and Control, Collector and Analyzer. The network configuration function module receives network configuration policies delivered by the application and management system, converts the policies into network device configuration for performance measurement, and delivers the instructions to network forwarding devices to enable the IFIT function. The collector and analyser receives and stores measurement data exported from network devices, then analyses and processes the data, such as fault location and performance deterioration alarm. At the same time, relevant measurement data and analysis results are reported to the application and management system.

c) An IFIT-enabled forwarding devices perform in-band flow quality measurement at the granularity of data packets in the IFIT domain. Based on the roles of the IFIT function, IFIT-enabled nodes (devices) are classified into the following roles:

   - The IFIT Head Node is responsible for adding an IFIT instruction header to a data packet of a specified flow object. The instruction header specifies the information to be measured in inband mode.

   - IFIT Transit Node, which identifies IFIT-enabled data flow packets, parses IFIT instruction header, and collects measurement data based on the IFIT instruction. Then the data collected in the transit node is stored in data packets or directly exported to the controller as required.

   - IFIT End Node identifies IFIT-enabled data flow packets, decapsulates IFIT headers, removes IFIT instruction headers, and extracts the quality measurement data carried in the data packet to the controller. Then end nodes forward the data packet.

d) The South Bound Interface (SBI), which is the interface used by the Controller to configure and collect telemetry data (e.g. OAM results, statistics, states, etc.) from the network nodes.

5.2 Closed-Loop Performance-Management Approach

This clause discusses relevant mechanisms to apply the Closed-Loop approach of the Reactive In-situ Flow Information Telemetry. In particular it is reported how this approach has been introduced in some relevant documents in IETF IPPM WG (e.g. IETF RFC 8321 [i.8] and IETF RFC 8889 [i.11]) to enable flexible and adaptive performance measurement.

IETF RFC 8321 [i.8] applies to point-to-point unicast flows and BUM traffic, while in general it is defined the Clustered Alternate-Marking method that is valid for multipoint-to-multipoint unicast flows, anycast and ECMP flows.

Therefore, the Alternate-Marking method can be extended to any kind of multipoint-to-multipoint paths, and the network-clustering approach is the formalization of how to implement this property and allow a flexible and optimized performance measurement support for network management in every situation.

Without network clustering, it is possible to apply Alternate Marking only for all the network or per single flow. Instead, with network clustering, it is possible to use the partition of the network into clusters at different levels in order to perform the needed degree of detail. In some circumstances, it is possible to monitor a multipoint network by analysing the network clustering, without examining in depth. In case of performance degradation, 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 IETF RFC 8321 [i.8].

This approach fits very well with the Closed-Loop Network and Software-Defined Network (SDN) paradigm, where the SDN orchestrator and the SDN controllers are the brains of the network and can manage flow control to the switches and routers and, in the same way, can calibrate the performance measurements depending on the desired accuracy. An SDN controller application can orchestrate how accurately the network performance monitoring is set up by applying the Multipoint Alternate Marking as described in the present document.

The monitoring network can be considered as a whole or split into clusters that are the smallest subnetworks (group-to-group segments), maintaining the packet-loss property for each subnetwork. The Network Clusters partition divides the Network Graph into the smallest subnetworks called Clusters. These Clusters can be combined and used at different levels to perform the needed degree of detail.

A possible algorithm for Cluster partition is a two-step algorithm (Iterative clustering algorithm):

1) Group the links where there is the same starting node.

2) Join the grouped links with at least one ending node in common.

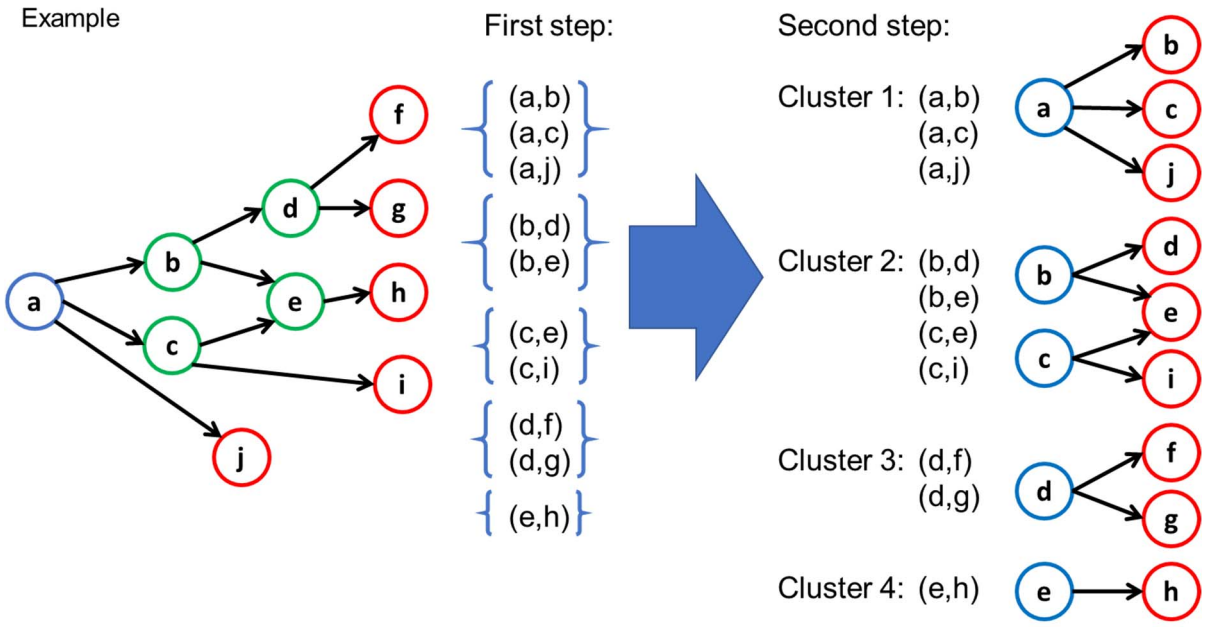

Figure 2: Example of the Iterative cluster algorithm for Cluster partition

The complete Alternate Marking framework is presented in IETF RFC 8889 [i.11].

Packet Loss can be measured on Cluster basis or by considering a combination of Clusters; and the borderline cases of single flows and whole network.

Delay measurements can be done in different ways:

- multipoint path basis measurement: the delay value is representative of an entire multipoint path (and clusters). The mean delay for a multipoint path can be defined;
- single packet basis measurement: the multipoint path is used just to easily couple packets between inputs and output nodes of a multipoint path. Hashing (IETF RFC 5475 [i.26]) and Multipoint Alternate Marking are coupled in this case. Clusters simplify the correlation of the hash samples from a topological point of view in terms of space, while Marking method anchor the samples to a specific period and simplify the correlation in terms of time.

By using these techniques, an SDN controller or a Network Management System (NMS) can calibrate performance measurements, since they are aware of the network topology. They can start without examining in depth. In case of necessity (packet loss is measured or the delay is too high), the filtering criteria could be immediately reconfigured in order to perform a partition of the network by using clusters and/or different combinations of clusters. In this way, the problem can be localized in a specific cluster or a single combination of clusters, and a more detailed analysis can be performed step by step by successive approximation up to a point-to-point flow detailed analysis. This is the so-called "closed loop".

This approach can be called "network zooming" and can be performed in two different ways:

1) change the traffic filter and select more detailed flows;

2) activate new measurement points by defining more specified clusters.

The next figures show a possible network to monitor and a possible application of the "network zooming" approach.

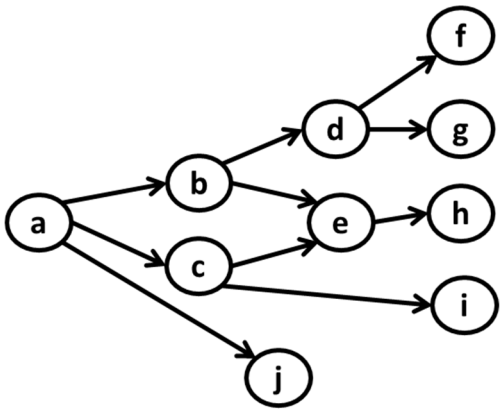

Figure 3: Example of network to monitor

In the beginning everything is good: Packet Loss = 0 and Delay/Jitter less than SLA values. The counters are activated only at the edge nodes.

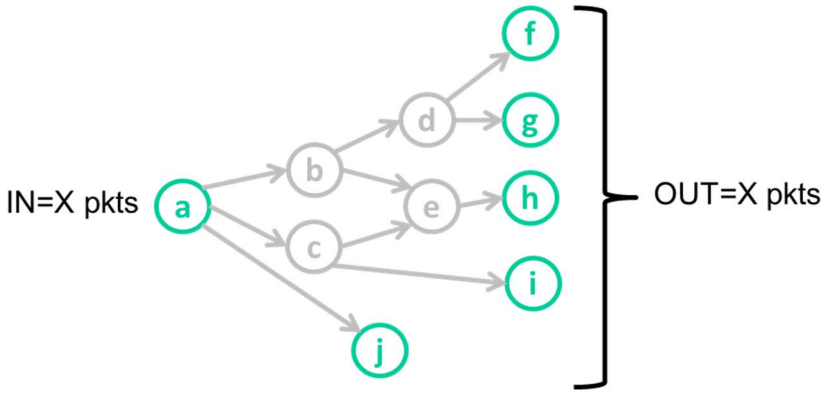

Figure 4: The full network is ok and is not experiencing losses

A Packet Loss event may be measured for the Full Network.

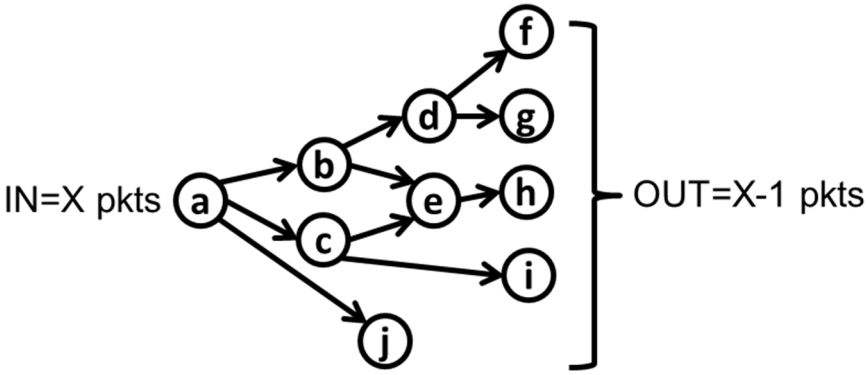

Figure 5: The full network is experiencing the loss of one packet

The next step is to configure Clusters Partition and locate which Cluster has the problem. The 4 Clusters are identified by applying the algorithm described before.

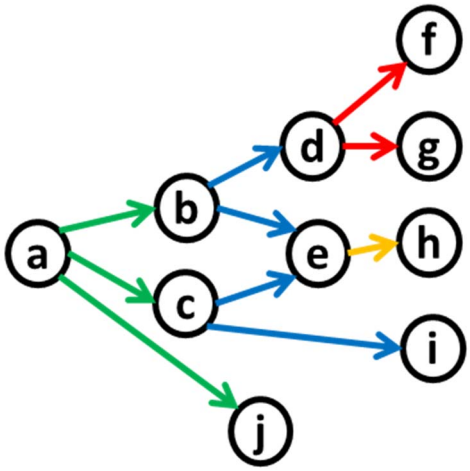

Figure 6: Cluster partition of the network

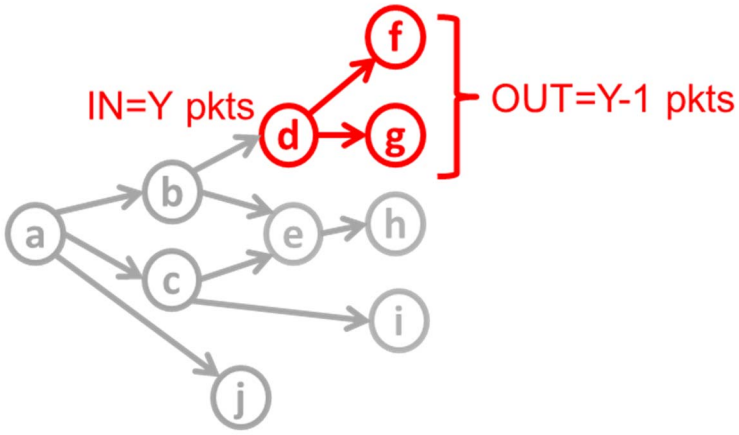

Figure 7: Identification of the Cluster where there is the loss

Finally, through more specific traffic filter, the fault is identified.

Figure 8: Identification of the link where the fault is present

Note that the network-zooming approach implies that some filters or rules are changed and that therefore there is a transient time to wait once the new network configuration takes effect. This time can be determined by the Network Orchestrator/Controller, based on the network conditions.

The Multipoint Alternate-Marking framework that is introduced in IETF RFC 8889 [i.11] adds flexibility to Performance Management (PM), because it can reduce the order of magnitude of the packet counters. This allows an SDN orchestrator to supervise, control, and manage PM in large networks.

The documents [i.12] and [i.13] define an architecture where the centralized Data Collector and Network Management can apply the intelligent and flexible Alternate-Marking algorithm as previously described.

As for IETF RFC 8321 [i.8], it is possible to classify the traffic and mark a portion of the total traffic. For each period, the packet rate and bandwidth are calculated from the number of packets. In this way, the network orchestrator becomes aware if the traffic rate surpasses limits. In addition, more precision can be obtained by reducing the marking period; indeed, some implementations use a marking period of 1 sec or less.

It is important to mention that the Multipoint Alternate Marking framework also helps Traffic Visualization. Indeed, this methodology is very useful for identifying which path or cluster is crossed by the flow.

5.3 Relationship with Network Telemetry Framework

Clause 5.4.3.3.1 in ETSI GR ENI 009 [i.14] introduces a Network Telemetry framework. This framework integrates multiple telemetry and data collection approaches, which allows flexible combinations for different telemetry data acquisition from different applications. As defined in ETSI GR ENI 009 [i.14], the components of the network telemetry framework include Data Source, Configuration, Telemetry Collector and Database, and Telemetry Data User.

IFIT-based Reactive Telemetry fits in the category of forwarding-plane telemetry and deals with the specific on-path technical branch of the forwarding-plane telemetry. The key functional components of IFIT-based reactive telemetry also match the components in Network Telemetry. "Intelligent Flow, Packet, and Data Selection" is responsible for realizing the quality measurement of specific flows/packets/data according to specific service requirements, matching the "Configuration" component. "Intelligent Data Export" is responsible to improve the transmission efficiency of collected information, matching the "Telemetry Collector and Database" component. "Dynamic Network Probe" is designed to flexibly obtain customized measurement information and improve the efficiency of network quality measurement, matching the "Data Source" component. "On-demand Underlying Technique Selection" is responsible to select various telemetry methods to realize different metrics measurement, matching the "Telemetry Data User" component.

6 Technical Requirements in IFIT-based Reactive Telemetry Framework

6.1 Key Components Overview

As shown in the IFIT-based reactive telemetry framework, the key components of IFIT are as follows:

1) Intelligent flow, packet, and data selection component realizes the quality measurement of specific flows/packets/data according to specific service requirements.

2) Intelligent data export component is based on de-redundancy and high-efficiency compression technology to improve the transmission efficiency of collected information.

3) Dynamic network probe can flexibly obtain customized measurement information and improve the efficiency of network quality measurement by deploying programmable hardware or software probes.

4) On-demand underlying technique selection component realizes different information measurement in different scenarios.

5) IFIT Network Automation mechanisms for the South Bound Interface (SBI), which is the interface used by the Controller to configure and collect telemetry data (e.g. OAM results, statistics, states, etc.) from the network nodes.

6.2 Intelligent Flow, Packet, and Data Selection

Network quality measurement such as IFIT will inevitably increase the consumption of network bandwidth, and cause an impact on forwarding performance. It is impractical to enable IFIT for all flows or packets in the network. Therefore, it is necessary to select some specific service flows, packets or data according to service or operation and maintenance requirements.

In the data plane, the Access Control List (ACL) provides a method to identify and select flows. For specific service, the sampling rate, packet measurement indicators, and collection nodes can be set as arguments to enable IFIT. According to different applications, any node can be allowed to receive or reject the collection of flows or packets. Based on these flexible mechanisms, IFIT can implement intelligent flow, packet and data selection and monitoring strategies to meet measurement requirements. In addition, IFIT can dynamically adjust selection and collection strategies in real time based on network load, forwarding processing capabilities, and other criteria.

Typical application scenarios for intelligent flow, packet and data selection include elephant flow recognition based on the sketch algorithm and adaptive packet sampling scenarios.

As elephant flow consumes large bandwidth and is sensitive to network changes, it has become the focus of performance measurement for network operators. By adopting the Count-Min Sketch algorithm in the IFIT-enabled node, the elephant flow can be periodically identified and reported to the controller. The controller generates a corresponding flow monitoring or measurement strategy based on the current elephant flow situation in the network, and sends it to the IFIT-enabled forwarding devices, thereby realizing the performance measurement of the elephant flow.

Applying IFIT to all packets on a specified stream may also over capacity. At this time, the measurement overhead can be reduced by adopting the packet sampling method on the specific flow. In the initial state, it is difficult to set an appropriate sampling frequency because the real-time bandwidth overhead information of the flow cannot be obtained. If the frequency is too high, it will consume large network resources, which may affect the network forwarding performance and even cause packet loss. Conversely, too low a frequency will result in loss of information and inaccurate measurements. For this scenario, the sampling frequency can be dynamically adjusted in real time based on network conditions. In order to avoid network congestion, the controller can collect relevant parameters to measure network congestion, such as packet delay, packet loss, etc. According to these collection messages, it is convenient to adjust the sampling frequency of IFIT measurement in real time, So as to achieve high-performance network quality measurement without affecting network forwarding performance.

6.3 Intelligent Data Export

IFIT can measure and export flow or packet quality information in real time. But there is a lot of redundancy in the collected information, and the high-density service flow quality measurement information uploading will consume a lot of bandwidth and may cause congestion of the exporting channel. Therefore, in order to reduce the transmission bandwidth and reduce the processing burden of the controller, it is necessary to perform de-redundancy and compression processing of the exported data.

Binary-based data transmission coding is an efficient method to export data, which can greatly reduce the amount of data transmission, such as Google Protocol Buffer (GPB) coding technology. In addition to effective data coding, IFIT can also collect information that does not time sensitive, cache them and send accumulated data in batches. In the process of batch data, a variety of redundant data deletion and compression technologies can be used. From the perspective of IFIT operation and maintenance, the caching batch data and exporting method is usually suitable for special error events. If the forwarding delay of the flow/packet exceeds the threshold and the forwarding path of the flow/packet is changed, there is no need to send all the original data to the controller, but only the relevant data before and after the changes.

A typical application scenario for realizing intelligent data export is that real-time monitoring data export scenario triggered by an abnormal event. Network operation and maintenance personnel often pay more attention to real-time and accurate perception of some network abnormal events, such as path changes, network congestion, and packet loss and so on. These abnormal events can be monitored through IFIT technology, such as encapsulating path tracking information in packets, and making time-stamps on network inbound and outbound interfaces. When a network device detects an abnormal event, it can describe the abnormal event through a strategy and send it to the controller. For example, when a flow has a forwarding node change, a path change event is triggered; when a packet is forwarded at a network node with a delay exceeding the delay threshold, a congestion event is triggered; when a packet is finally discarded due to buffer overflow, the packet loss event is triggered. Through abnormal event monitoring, network forwarding nodes only need to export abnormal events and related quality measurement information to the controller, which can greatly reduce the amount of data export.

In addition, IFIT can also use the general IP data export technology (i.e. IPFIX) to realize the export of measurement data. IPFIX is a template format-based information export protocol based on data feature analysis. It can obtain different data formats based on different collection requirements with strong scalability.

6.4 Dynamic Network Probe

Limited by data plane resources, it is difficult to achieve comprehensive monitoring of network data. On the one hand, hardware resources such as Network Processor (NP) and Traffic Manager (TM) components in device are the key factors to realize high-performance forwarding. At the same time, a large amount of bandwidth resources are consumed in order to process and forward massive messages. Therefore, hardware resources in equipment and network bandwidth resources become scarce resources in the network. On the other hand, intelligent applications also have diversity and real-time variability requirements for measurement data. Therefore, it is very important to meet the dynamic data measurement requirements under limited resource conditions.

Data plane programmability allows IFIT to dynamically load new data probes, namely Dynamic Network Probes (DNP). DNP is a technology that enables probes for customized data collection in different network planes, and can be loaded into the data plane through incremental programming or configuration. DNP can effectively perform data generation, processing and aggregation, and introduces sufficient flexibility and scalability for IFIT. According to DNP technology adopted in IFIT, through on-demand detection, not only can the optimization of data export be realized, but also the customization of detection information can be realized based on service requirements.

6.5 On-demand Underlying Technique Selection

IFIT is a set of technologies including data collection and export techniques, so it can flexibly adapt to different network conditions and different application requirements. For example, depending on the types of data that are interest, IFIT may choose either IOAM or PBT to collect the data. if an application needs to track down where the packets are lost, switching from IOAM to PBT should be supported.

IFIT can further integrate multiple data plane monitoring and measurement techniques together and present a comprehensive data plane telemetry solution. Based on the application requirements and the real-time telemetry data analysis results, new configurations and actions can be deployed.

6.6 IFIT Network Automation

This clause discusses the existing and proposed mechanisms for the South Bound Interface (SBI), which is the interface used by the Controller to configure and collect telemetry data (e.g. OAM results, statistics, states, etc.) from the network nodes. The North Bound Interface (NBI) is the interface between the Service Orchestrator and the Controllers.

The flexibility and dynamicity of the IFIT applications are given by the use of additional functions on the controller and on the network nodes, and this can be done by adding a telemetry information exchange between the network nodes and the controllers in order to enable the so-called Closed-Loop automation.

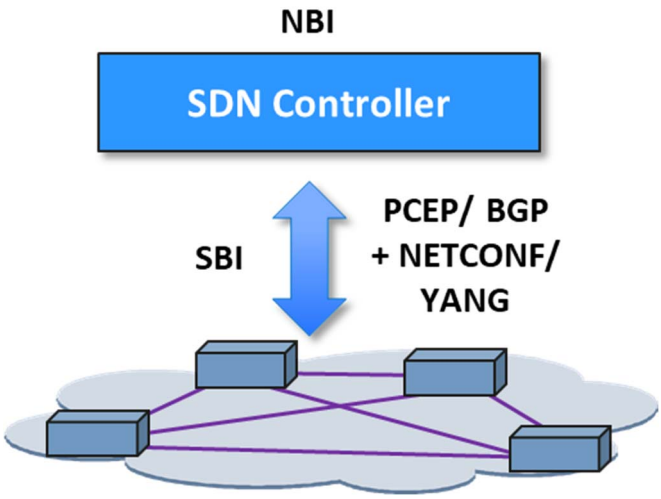

Figure 9: SDN Controller-Network interfaces

In this regard it is possible to mention the Model Driven Telemetry (MDT) that enables the Closed Loop Automation. MDT is a new approach for network monitoring in which data is streamed from network devices continuously using a push model and provides near real-time access to operational statistics. The configuration is done with Data Models and Telemetry is also done with Data Models. Model Driven Telemetry is also known as YANG Push and defined in IETF RFC 8639 [i.15], IETF RFC 8640 [i.16], IETF RFC 8641 [i.17] and IETF RFC 8650 [i.18]. Applications can subscribe to specific data items they need, by using standard-based YANG data models over NETCONF YANG.

The YANG module specified in [i.19] defines a data model for IOAM capabilities using the YANG data modelling language (see IETF RFC 7950 [i.20]). It is designed to be used by the network management protocols such as NETCONF [i.21] or RESTCONF [i.22] in order to configure the network nodes. It supports all the five IOAM options, which are Incremental Tracing Option, Pre-allocated Tracing Option, Direct Export Option, Proof of Transit Option, and Edge-to-Edge Option. IOAM can surely leverage YANG Push to achieve flexible telemetry.

In addition to YANG models, other protocols are used for the communication between the control layer and the network nodes: Path Computation Element communication Protocol (PCEP) and Border Gateway Protocol (BGP).

An automatic network requires the Service Level Agreement (SLA) monitoring on the deployed service. So that the system can quickly detect the SLA violation or the performance degradation, hence to change the service deployment. In this regard, [i.23] and [i.24] define extensions to BGP and PCEP respectively in order to distribute IFIT information. So that IFIT behaviour can be enabled automatically when the path is instantiated.

The definition of the IFIT data plane methods for SR-MPLS and SRv6 imply requirements for various routing protocols, such as BGP and PCEP. [i.23] aims to define BGP extensions to distribute SR policies carrying IFIT information and this allows to signal the IFIT capabilities in order to automatically configure and run IFIT methods when the SR Policy candidate paths are distributed through BGP. Similarly, the PCEP extension defined in [i.24] allows to signal the IFIT capabilities and apply the IFIT attributes for all path types, as long as they support the relevant data plane telemetry method. In this way IFIT methods are automatically activated and running when the path is instantiated.

In summary, by combining the use of YANG Push, PCEP and BGP it is possible to obtain the reactive and adaptive telemetry for IFIT methodologies.

7 Examples of applications and scenarios

7.1 Generic description of application and scenarios

This clause reports both real and experimental applications where the Reactive IFIT can be applied and the related benefits of flexible and adaptive performance measurement are analysed.

There are application fields where it may be useful to take into consideration the Multipoint Alternate Marking IETF RFC 8889 [i.11]:

- VPN: The IP traffic is selected on IP source basis in both directions. At the endpoint WAN interface all the output traffic is counted in a single flow. The input traffic is composed by all the other flows aggregated for source address. So, by considering n end-points, the monitored flows are n (each flow with 1 ingress point and (n-1) egress points) instead of n*(n-1) flows (each flow, with 1 ingress point and 1 egress point);

- Mobile Backhaul: LTE traffic is selected, in the Up direction, by the EnodeB source address and, in Down direction, by the EnodeB destination address because the packets are sent from the Mobile Packet Core to the EnodeB. So the monitored flow is only one per EnodeB in both directions;

- Over The Top (OTT) services: The traffic is selected, in the Down direction by the source addresses of the packets sent by OTT Servers. In the opposite direction (Up) by the destination IP addresses of the same Servers. So the monitoring is based on a single flow per OTT Servers in both directions.

- Enterprise SD-WAN: SD-WAN allows to connect remote branch offices to Data Centres and build higher-performance WANs. A centralized controller is used to set policies and prioritize traffic. The SD-WAN takes into account these policies and the availability of network bandwidth to route traffic. This helps ensure that application performance meets Service Level Agreements (SLAs). This methodology can also help the path selection for the WAN connection based on per Cluster and per flow performance.

Note that the list is just an example and it is not exhaustive. More applications are possible.

7.2 Performance Measurement and Fault Isolation in 5G Transport Network

The 5G transport network has various access modes and carries various mobile transport services (such as HD video) that pose higher requirements on link connectivity and performance metrics (e.g. packet loss, latency, jitter). To ensure high-quality, stable, and reliable network services for the 5G transport network, it is an effective measure to deploy IFIT performance monitoring for N2/N3 (SCTP/GTP) traffic.

In the 5G transport network, the flow detection feature provided by IFIT can be used to quickly demarcate and locate network faults, thus improving OAM efficiency. When the controller finds that the SLA does not meet the service requirements through IFIT End-to-End measurement, it will automatically perform the IFIT trace measurement to demarcate poor quality point hop-by-hop. The hop-by-hop detection results can be checked through the application management system.

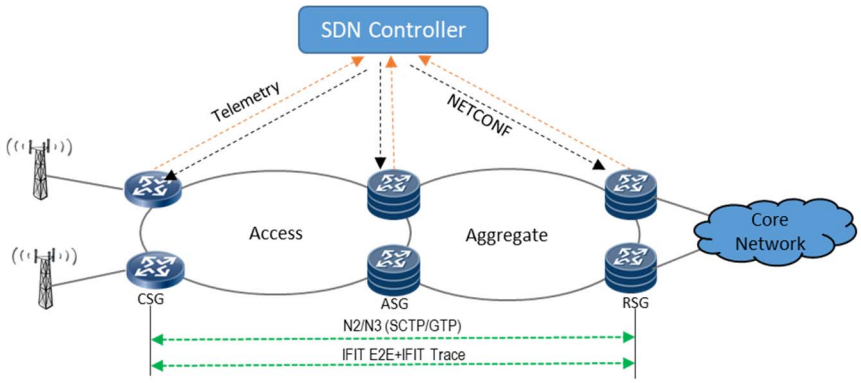

Figure 10: Application of IFIT in 5G Transport Network

As shown in Figure 10, the SCTP/GTP traffic are transmitted between the base station and AMF/UPF through N2/N3 interfaces respectively. In this scenario, both IFIT E2E and IFIT Trace measurement solutions are implemented in the 5G transport network.

Firstly, IFIT E2E measurement is performed, which checks whether the E2E monitoring of the N2/N3 (SCTP/GTP) traffic of the base station is abnormal. If the E2E measurement result is normal, the fault on the bearer network is preliminarily excluded. If the E2E measurement result is abnormal, the IFIT trace measurement.

Then, IFIT trace (Hop-by-Hop) measurement is triggered when the flow performance indicator exceeds the specified threshold. In this case, SDN controller summarizes the reported IFIT trace measurement data for fault locating. For example, if the packet loss rate or delay between ASG and RSG exceeds the threshold, then the fault point is quickly located. And based on this, the further root cause can be analysis and concluded.

According to the traffic monitoring and fault detection, IFIT technology provides a way to accurately measure service quality for real service flows, and provides a quickly fault location means by hop-by-hop detection, thus greatly improving the network OAM efficiency.

7.3 IFIT-based Reactive Telemetry Loop within ENI System

As specified in ETSI GS ENI 005 [i.25], clause 6.3.1.4 and clause 6.3.1.5, there are five functions (i.e. Knowledge Management, Context Awareness, Cognition Management, Situational Awareness, Model-Driven Engineering and Policy Management) are represented by one or more ENI Functional Blocks ENI Functional Architecture with Control Loops and Domains.

Meanwhile, External Reference Points (see clauses 4.4.6.1, 7.2 and 7.3 in ETSI GS ENI 005 [i.25]), and Internal Reference Points (see clauses 4.4.6.2, 7.6 and 7.7 in ETSI GS ENI 005 [i.25]) are used by ENI System to communicate with the Assisted System (or its Designated Entity) and communicate between different ENI System Functional Blocks respectively.

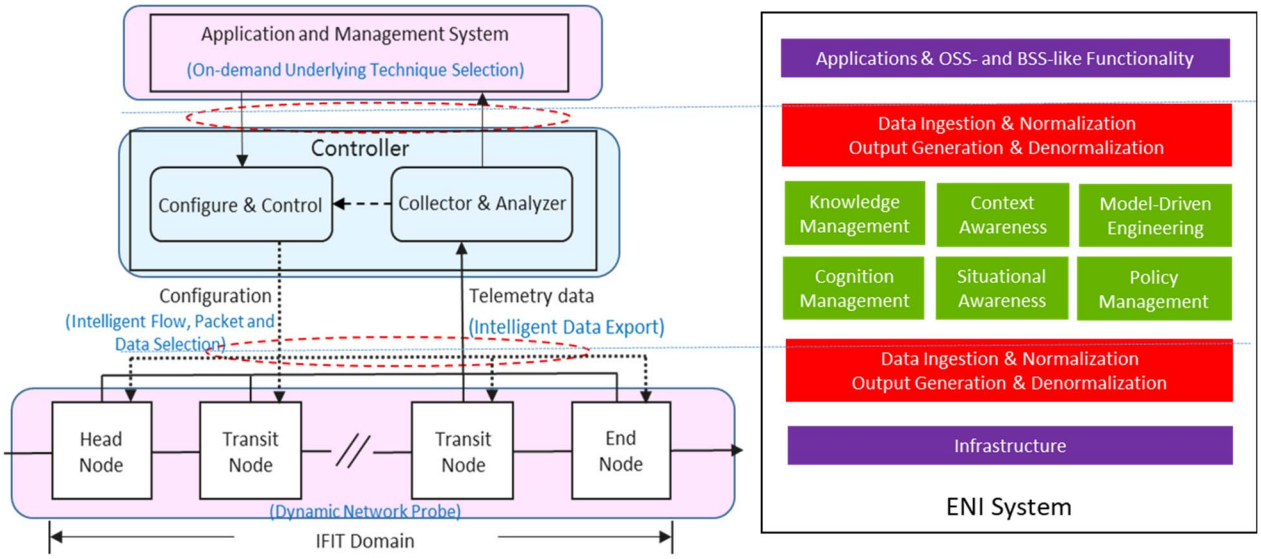

Figure 11: An Exemplary IFIT-based Reactive Telemetry Loop within ENI System

As shown in Figure 11, mapping these functional components of IFIT-based Reactive Telemetry Framework described in clause 5.1 into the ENI Functional Architecture with Domains and Control Loops, it can be found that:

a) The Application and Management System and the IFIT-enabled forwarding devices work as Inputs and Outputs to ENI, which locate in Domain 1. On the one hand, the IFIT network configuration policy generated by the application and management system, which includes information such as a specified flow object to be measured, a performance indicator to be collected, and a test data exporting mode (passport mode or postcard mode), are transmitted to the API Broker, which then communicates the data using one (or more) of the designated ten input External Reference Points. Each input first goes to the Data Ingestion and then to the Normalization Functional Blocks. At this point, the data is in a format that can be understood by the six Internal ENI Functional Blocks. On the other hand, the application and management system receives IFIT quality measurement data and analysis results from the collector and analyser, which are translated into specific formats required by the application and management system through Output Generation Functional Block, then displays the results in a visualized manner.

b) The Controller consists of both Collector and Analyzer and Configuration and Control components locate in Domain 2 and Domain 3 respectively, which are realized in related ENI Architectural Functional Blocks. On one hand, the collector and analyser receive and store measurement data exported from network devices, then analyses and processes the data, such as fault location and performance deterioration alarm, which is realized within Knowledge Management Functional Block, Context-Aware Management Functional Block, Situational Awareness Functional Block and Model Driven Engineering Functional Block. At the same time, relevant measurement data and analysis results are reported to the application and management system. On another hand, the network configuration function module receives network configuration policies delivered by the application and management system, converts the policies into network device configuration for performance measurement, and delivers the instructions to network forwarding devices to enable the IFIT function, which are realized within Cognition Management Functional Block, Model Driven Engineering Functional Block and Model Driven Engineering Functional Block.

c) The IFIT-enabled forwarding devices work as Inputs and Outputs to ENI, which locate in Domain 1. An IFIT-enabled forwarding devices perform in-band flow quality measurement at the granularity of data packets in the IFIT domain. Similarly, performance measurements metrics are transmitted to the API Broker, which then communicates the data using one (or more) of the designated ten input External Reference Points. Each input first goes to the Data Ingestion and then to the Normalization Functional Blocks. At this point, the data is in a format that can be understood by the six Internal ENI Functional Blocks.

8 Conclusions and Recommendations

The present document describes some new technical methods to meet the requirements for improving traditional network OAM methods and meet users' requirements for E2E high-quality network experience in data-driven intelligent networks. IFIT-based reactive telemetry works with telemetry, big data analytics, intelligent algorithms, and other technologies to build an intelligent closed-loop OAM system. The use of IFIT provides data sources for the big data platform and intelligent algorithm analysis. These results form the foundation of the intelligent OAM system's ability to implement e.g. precise fault demarcation and locating as well as fast fault self-healing.

IFIT methodologies leverage the ENI architecture (ETSI GS ENI 005 [i.25]) in order to allow a reactive performance management. The Functional Block-based architecture helps to address the application of an intelligent, flexible and adaptive approach with IFIT techniques in order to meet the network conditions and user needs.

Hence, the main conclusion of the present document is that IFIT-based reactive telemetry is highly recommended to be used as one kind of data source and telemetry processing in the Data Ingestion Functional Block specified in ETSI GS ENI 005 [i.25]. In particular, the present document enriches the Data Ingestion Functional Block by describing key technical components for telemetry processing including Intelligent flow, packet, and data selection, Intelligent data export, Dynamic network probe, and On-demand underlying technique selection. The results of the investigation could be applied to the normative work for ETSI GS ENI 005 [i.25].

Furthermore, a synergy between IETF, that is responsible for the definition of IFIT related methodologies, and ETSI, that is working towards achieving an ENI architecture, is also expected.

Annex A:
Change History

Date	Version	Information about changes
2020-06	V0.0.1	Initial early draft with skeleton
2020-09	V0.0.2	New baseline with ToC updates and Reactive Telemetry Framework inputs
2020-10	V0.0.3	New baseline with clause 5.2, clause 7 and Annex Author & contributors
2021-03	V0.0.4	New baseline with clause 4.1, clause 6, clause 7.1 and Annex Author & contributors
2021-06	V0.0.5	New baseline with clause 2.2, clause 5.3, clause 6.6 inputs
2021-07	V0.0.6	New baseline with clause 7.2, clause 8, clause 5.2 and clause 6.6 changes
2021-08	V0.0.7	New baseline with clause 3.2, clause 7.2 and clause 7.3 changes
2021-09	V0.0.8	New baseline with no editor's notes and revised References and abbreviations
2021-12	V0.0.12	New baseline with clause 4.1 and clause 8 changes
2021-12	V0.0.13	New baseline with informative references

History

Document history		
V1.1.1	March 2022	Publication