

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: January 14, 2021

A. Karaagac
J. Hoebeke
Ghent University - imec
July 13, 2020

In-band Network Telemetry for 6TiSCH Networks
draft-karaagac-6tisch-int-01

Abstract

This document describes In-band Network Telemetry for 6TiSCH Networks, offering a flexible monitoring solution with minimal resource consumption and communication overhead while supporting a wide range of monitoring operations and strategies for dealing with various network scenarios and use cases. It enables 6TiSCH networks to collect per-packet and per-hop monitoring information by piggybacking telemetry information onto the data packets by exploiting the remaining space in the IEEE 802.15.4e frames, thus not impacting network behavior and performance. This document also discusses the data fields and associated data types for 6TiSCH INT mechanism.

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 January 14, 2021.

Copyright Notice

Copyright (c) 2020 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. Terminology	4
2. In-band Network Telemetry for 6TiSCH	4
2.1. Capacity-Neutral Network Monitoring	5
2.2. INT Data Model, Format and Encoding	6
2.2.1. INT Sub-IE Format	8
2.2.2. Telemetry Data Model	10
2.3. INT Strategies	12
2.3.1. Opportunistic Logic	12
2.3.2. Probabilistic Logic	12
3. Acknowledgements	13
4. IANA Considerations	13
4.1. IETF IE Subtype INT	13
5. Security Considerations	13
6. References	13
6.1. Normative References	13
6.2. Informative References	14
6.3. External Informative References	14
Authors' Addresses	15

1. Introduction

For continuous, persistent and problem-free operation of "IPv6 over the TSCH mode of IEEE 802.15.4e" (6TiSCH) Networks [I-D.ietf-6tisch-architecture], it is critical to have visibility and awareness into what is happening on the network at any one time. For centrally managed 6TiSCH networks, it is required to collect and analyze network performance data, often as close to real time as possible. For TiSCH networks with distributed management solutions, it is still vital to monitor network nodes continuously or periodically to ensure their functioning, detect relevant problems, perform traffic engineering and network optimization.

Nevertheless, efficient monitoring and management mechanisms for these networks have not been addressed adequately. First, traditional active network and health monitoring systems (i.e. statistical polling, active probing) are of limited applicability in these constrained and dynamic networks due to their static and inefficient design. Especially, considering the constrained nature

of sensor networks, the introduced control traffic can occupy extensive network resources, impact network behavior and/or interfere with the scheduled application traffic flow. Secondly, the passive health monitoring and tomography methods can only offer limited capabilities for collecting in-network state information and telemetry data, thus are not sufficient for advanced network monitoring and fine-grained management operations. In addition, the 6TiSCH WG is defining a management interface, based on CoAP Management Interface (CoMI) [I-D.ietf-core-comi], which can be used to monitor network performance and perform network configurations [I-D.ietf-6tisch-coap]. However, performing telemetry via CoMI interfaces will result in a polling-based monitoring scheme which may cause a large amount of control traffic.

This document specifies an In-Band Network Telemetry (INT) mechanism adapted to 6TiSCH Networks. It provides the definition of telemetry semantics and data models for 6TiSCH Networks and their efficient encoding in the IEEE 802.15.4e [IEEE802154e] frames. Additionally, it defines a set of novel telemetry operations and strategies for dealing with various network scenarios and system interactions.

The proposed INT-based network monitoring solution creates an efficient, adaptive and flexible design which offers several novel monitoring functionalities and telemetry operations for 6TiSCH Networks.

- o Opportunistic piggybacking mechanism that eliminates the need for artificial probing packets and resource reservation for monitoring data in 6TiSCH Networks.
- o Real-time monitoring capabilities where the collected telemetry data reflects the momentary network performance and the exact treatment that an application packet encounters.
- o The combination of real-time edge-to-edge packet-level network information (e.g. reliability, latency) and hop-by-hop telemetry data (e.g. per-hop latencies, queue states and link qualities).
- o Flexibility in terms of telemetry initiation and addition approaches: continuous, periodic, event-driven or query-driven.
- o Flexibility for forwarding nodes to initiate an INT operation on a packet with another source.
- o Flexibility for source and forwarding nodes to decide what to add: even a subset of INT entries if not all of them fit in the frame.

1.1. Terminology

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Readers are expected to be familiar with terms and concepts defined in [IEEE802154e] and [I-D.ietf-6tisch-architecture].

"RPL", "RPL Dag Rank", MaxRankIncrease, MinRankIncrease and RootRank are defined in the "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks" [RFC6550] specification.

This document refers also to the following terminology.

INT : In-band Network Telemetry

E2E : End-to-End

HBH : Hop-by-Hop

2. In-band Network Telemetry for 6TiSCH

INT, or also referred to as In-situ Operations, Administration, and Maintenance (iOAM) [I-D.ietf-ippm-ioam-data], is created to complement current out-of-band monitoring mechanisms and allows for telemetry metadata to be collected as packets traverse a network. The term "in-band" refers to the fact that telemetry data is carried within data packets rather than being sent within specifically dedicated packets. Therefore, it does not require artificial probing packets or dedicated middle-boxes, and the network state is obtained at the exact point in time the real user traffic passes through. Also, the insertion of in-band information does not change the forwarding behavior of the packet. However, it might impact the packet delivery ratios (PDR) due to the increase in the length of the transmitted frames.

The 6TiSCH INT mechanism collects the telemetry data while a packet is traversing towards the Backbone Router. This measurement data can typically be node or network state information such as health/failure reports, link/neighbor statistics, network topology and node/link occupancy. When the packets reach the edge (backbone router) of the network, the telemetry metadata is removed and telemetry reports are generated to be used by the Network Management Entity (NME) for further visualization, analysis and management.

2.1. Capacity-Neutral Network Monitoring

In a Timeslotted Channel Hopping (TSCH) [RFC7554] network, time is globally synchronized and is sliced up into time slots. The time synchronization in the network means that all nodes share a timeslot counter, named Absolute Slot Number (ASN), indicating the total number of slots which have passed since the network has started [IEEE802154e]. The overall communication is orchestrated by a schedule which instructs each node what to do (transmit, receive, sleep) in each timeslot [IEEE802154e]. In this TSCH schedule, a single element, named cell, is identified by a pair of slotOffset and channelOffset, which is used to define the communication time and frequency.

The duration of a time slot is not defined by the standard, but it is defined to be long enough to send a data frame, handle the radio turnaround and receive an ACK, typically being 10ms. With radios that are compliant with IEEE 802.15.4 operating in the 2.4 GHz frequency band, a maximum-length frame of 127 bytes is considered, which takes around 4 ms to transmit [RFC7554]. Whatever size that a node is sending, the resources are reserved for that node so that it can transmit a data frame of 127 bytes. If the node has a shorter frame to send, there will be remaining time for that node to sleep or stay idle. That means the reserved time/bandwidth resources are wasted, instead of being used for other good reasons. Therefore, this paper proposes a mechanism that collects the monitoring information for each node by piggybacking telemetry information on the data packets in order to leverage these remaining resources, as presented in Figure 1. If there is no or insufficient remaining space in the transmitted frame, the node cannot add any telemetry information.

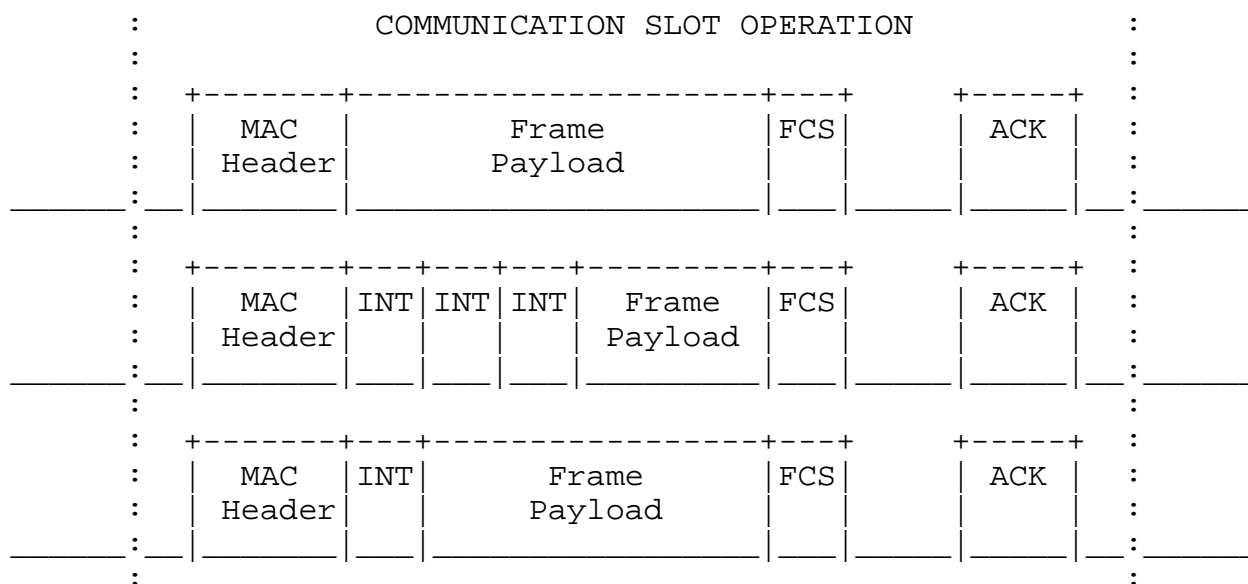


Figure 1: Capacity-Neutral Network Monitoring.

Regarding the cost of the INT operation, there will only be a limited amount of extra energy consumption for the transmitting and receiving nodes in order to transmit/receive extra bytes in the frame. However, it will not use any resource (i.e. slot, bandwidth) reserved for other application or control traffic and it will not have any effect on the network capacity, network behavior and traffic flows.

2.2. INT Data Model, Format and Encoding

For the insertion of telemetry data in IEEE 802.15.4 MAC frames, the Information Elements (IEs) are used, which are positioned between the end of the MAC Header and the Frame Payload. The IEs are intended to extend 802.15.4 in an interoperable manner and they can be exchanged between one-hop neighbors or forwarded for communication towards further away devices, thus allowing several optimizations [IEEE802154]. The IEs are structured containers as Type, Length, Value fields (TLV) and they have two types, named Header IEs and Payload IEs [IEEE802154]. Header IEs are part of the MAC header and most of their processing is done by the MAC, so IETF protocols should not have any direct effect on that processing. Contrary, Payload IEs are part of the MAC payload and they may be encrypted and authenticated. According to the standard, each frame can include one or more Header or Payload IEs that contain information.

IETF has formulated a request towards the IEEE 802.15 Assigned Numbers Authority (ANA) to allocate a registry number and described how IETF IEs should be formatted with their subtypes [RFC8137]. Also, 6TiSCH WG has expressed the need for IEs and a temporary

assignment is already provided [RFC8137]. For the design of IEs for INT data, an IETF INT sub-IE type is created by following the IETF IE subtype format.

For inserting an INT sub-IE in a MAC frame, the node first must set the "Information Elements Present" field in the 802.15.4 header. Next, Header IEs must be added which will be terminated by a Header Termination 1 IE (2 Bytes). If there is no Header IE, the Header Termination 1 IE must still be added in order to indicate the start of Payload IEs [IEEE802154]. After that, the IETF IE descriptor (2 Bytes: type, id, length) must be added, where the IETF IE Group ID is assigned as 0x5 in IEEE 802.15 ANA [ana2019]. Then, the INT sub-IEs must be added including the INT sub-IE descriptors (1 Byte: sub-IE ID) and the relevant INT data. At the end of the payload IEs, a Payload Termination IE (2 Byte) must be added. Considering all these necessary IEs, 7 Bytes of overhead will be added to the frame in order to insert any size of INT data. The resulting frame format after the INT sub-IE insertion is provided in Figure 2.

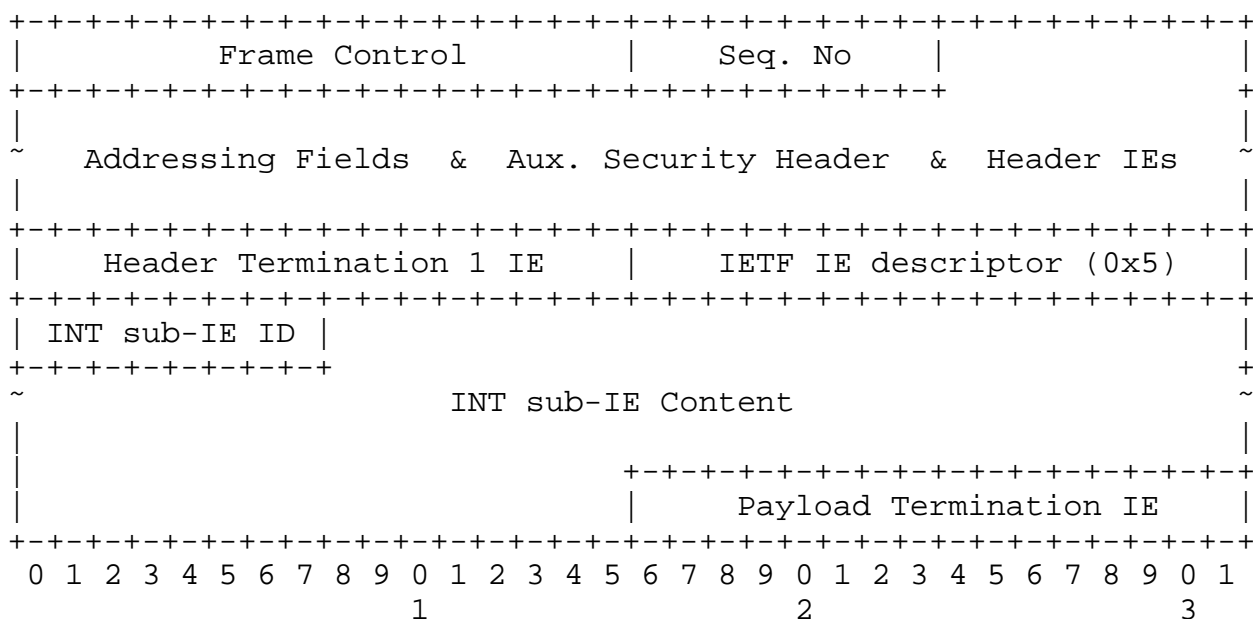


Figure 2: The frame format with inserted INT IE.

The following subsections describe the approach and format for embedding telemetry information in the body of an active data packet via IETF INT sub-IEs.

2.2.1. INT Sub-IE Format

The INT-extended packets in transit must contain telemetry instructions, so the network nodes can process and insert relevant telemetry data according to these instructions when processing the packets. In this regard, based on the requirements and targeted telemetry functionalities for 6TiSCH networks, the INT sub-IE format is designed with its headers and content, as shown in Figure 3. In this format, the Subtype Id represents the IETF IEs subtype identifier as defined in [RFC8137]: IANA_IETF_IE_INT.

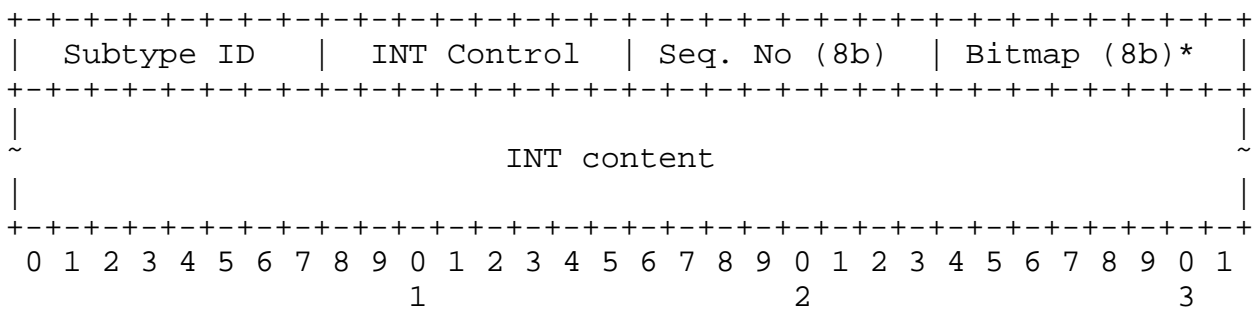


Figure 3: The format of the IETF INT IE Subtype.

The INT Header consists of three parts; INT Control header, Sequence Number and Bitmap. The INT Control header will be used to instruct the other nodes about the telemetry modes and functions considered in the particular packet. The detailed format of this field is provided in Figure 4. The sequence number is an 8-bit counter for the INT source, in order to differentiate between different INT data entries from the same node and to detect the end-to-end delivery ratio for data packets with INT entries. Finally, the Bitmap is the optional INT request vector where each bit represents another type of INT data. It is used to inform middle nodes about the relevant telemetry data to add or determine the content of the INT metadata during the decoding. The details of the INT control header is provided in the remainder of this subsection.

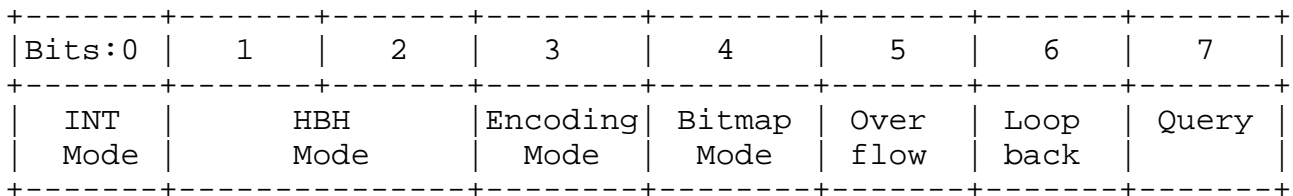


Figure 4: The format of the INT Control Field.

INT Mode (1b): defines the mode of telemetry operation: End-to-End (E2E) or Hop-by-Hop (HBH). In E2E mode, the middle nodes may

only forward the INT data without any processing or addition. This mode may be used to monitor end-to-end network performance or notify a central entity about local performance issues. On the other hand, HBH mode may be used to perform per-hop telemetry operation which allows all or a subset of the traversing nodes to add telemetry data if any space is left in the current frame.

HBH Mode (2b): defines the behavior of middle nodes in Hop-by-Hop telemetry operations. It must be 0 if End-to-End INT Mode is selected. If Mode 1 (Opportunistic) is selected, then all the nodes will try to add telemetry data in a opportunistic manner. Mode 2 (Probabilistic) will trigger the middle nodes to follow a probabilistic approach for telemetry addition. So the nodes may add telemetry data with a certain probability which can dynamically change based on the last time it added a telemetry, the available space in the forwarded frame and the remaining number of hops. This approach can be beneficial when attempts to add INT data frequently lead to frame size overflows and can enable collecting data from a more diverse set of nodes in the network. Finally, Mode 3 enables middle nodes to decide to add or skip telemetry data in distributed manner. In this mode, the nodes which detect performance drops/issues may add telemetry data to packets as a middle node. This will also avoid the usage of resources for already known/not important data.

Encoding Mode (1b): determines the encoding mode that will be used in the INT content. The first option is using the Bitmap mode (Content or Node) which must be followed by telemetry data as byte array. The type of each data will determine the length of that field which will be used to process/decode the data. The second option is using a TLV encoding, where each entry must be encoded with its type, length and value. This will bring flexibility to insert data with variable length and enable nodes to decide on the INT content to insert. In order to reveal the owner of each INT entry, each node must add a Node Id entry before the other telemetry data. In addition, the whole INT content should be processed to understand what kind of telemetry data is added by each node.

Bitmap Mode (1b): defines what kind of bitmap will be used: Content Bitmap vs Node Bitmap. If it is Content Bitmap, then that bitmap will apply for each node that adds INT data. Each node must follow the given bitmap and concatenate the relevant entries to the end of the current INT content. The content must include all fields mentioned in the bitmap with correct sizes. So, the bitmap can be used to detect the length of each field during decoding. Alternatively, the Node Bitmap option enables each node to add its own bitmap along with the INT data which will

bring independence to nodes for adding different kinds of INT data. During decoding, each node bitmap can be used to detect the length of each field.

Overflow (1b): states if any INT entry overflow has happened until that particular hop. If it is set, all of the following hops will know that they won't be able to add any INT entry, and so they can avoid any kind of INT processing.

Loopback (1b): may be used by the central entity to achieve downlink INT operation towards an end node. The central entity may insert an INT sub-IE entry with enabled loopback and then middle nodes may add INT data until it arrives at the destination node. After that, that node must forward the collected INT data to the central management entity in any of the following uplink data messages as INT entry. This downlink INT operation will still happen fully in-band.

Query (1b): may be used by the central unit to trigger an uplink INT operation with given configuration. When a node receives a packet with attached INT sub-IE including Query bit set, then it should create an INT operation using the received bitmap. This can be used to create a polling-based INT operation triggered by central entity. For instance, there can be the case that the central management entity detects a problem in the network, but there is not sufficient data to troubleshoot or isolate it. Then it can send a query to certain nodes to collect more insight about the problem.

2.2.2. Telemetry Data Model

Based on a number of monitoring and management scenarios for 6TiSCH Networks, a number of Telemetry Data types are defined. The proposed telemetry data model with limited scope is provided in Table 1 with details about their bitmap id, name, size and description. One can extend the INT Metadata by defining any relevant telemetry data types in order to collect other network status information; such as link quality, number of neighbors, number of incoming/outgoing cells, number of re-transmissions. As it is shown in Table 1, four of the bitmap ids are reserved for any further type definition.

Bitmap ID	Name	Size	Description
0	Node ID	2B	Device identifier (e.g. 802.15.4 16bit short address)
1	Receive Channel & Timestamp	2B	Channel (4b) & Reception or Generation time (12b)
2	Utilization indicator	1B	Transit Delay (4b), Queue Depth (4b)
3	RSSI	1B	Received Signal Strength (-127...0...127)
4-7	Reserved	-	Reserved for other telemetry data

Table 1: Telemetry Data Model

Node Id is one of the fundamental telemetry information types and represents the unique identifier of the node that inserts the telemetry data. In the scope of 6TiSCH networks, IEEE 802.15.4 16-bit short addresses can be used.

Receive Channel and Timestamp constitute a combined telemetry entry. The first 4 bits of this field represent the channel (0...15) the packet is received on, i.e. one of the available 16 IEEE 802.15.4 channels. The Timestamp represents the 12 least significant bits of the time (expressed in ASN which is 5 bytes) at which a packet that needs to be forwarded is received. For the source node, this time represents the time the packet is generated. Since all of the network nodes share the same ASN, the timestamps on each node are inherently synchronized. Assuming a 10ms slot length, 12 bits are enough to represent 40.96 seconds which is sufficient to detect all of the timestamps based on the reception ASN at the border router. The packet generation time can also allow us to understand the age of telemetry data and evaluate its validity.

Utilization indicator illustrates the node occupation when the packet traverses that node. The first 4 bits of this field represent the transit delay which is the delay (in slots) between the reception of a frame and its entry to the outgoing queue to be transmitted to the next hop. For the source node, this field will be 0. The remaining 4 bits constitute the Queue Depth value which is the number of packets in the outgoing queue at the time.

RSSI represents the received signal strength for that frame measured at the particular hop. It must take values between -127 dBm and 127

dBm. This value is 0 for the source node and will be ignored during INT processing.

2.3. INT Strategies

During INT entry initiation and addition process, the nodes can follow various INT strategies via making use of several locally calculated indicators. For instance, the nodes may avoid adding repetitive INT entries by checking the last time a similar INT operation is performed. Additionally, they may continuously process all locally collected telemetry data, detect events/misbehavior and assign an importance/relevance metric to each of them, then trigger an INT operation respectively.

2.3.1. Opportunistic Logic

In this strategy, each node tries to exploit immediate telemetry insertion opportunities, regardless of any planning or principle, in a greedy manner. So, the nodes will take every chance to insert telemetry in any suitable outgoing packet towards the border router.

Although this approach will maximize the total amount of collected telemetry, the source node and the nodes which are closer to the source will have a higher chance to insert telemetry data and subsequent nodes may not even get any chance to add any telemetry. This results in an unfair telemetry distribution and different INT inter-arrival times for different nodes. Therefore, for certain network scenarios, especially for large networks with limited telemetry opportunities, this approach may result in an inadequate network view due to the telemetry information that comes from only a limited part of the network.

2.3.2. Probabilistic Logic

In this strategy, the nodes are following a probabilistic approach where each node may insert or skip INT entries with certain probabilities which must be dynamically calculated in a distributed manner. This probability can be calculated based on the current frame size (including headers, payload, current INT), the size of a newly to be added INT entry based on Bitmap, and the remaining hop count that can be calculated based on the RPL Dag Rank and RPL link parameters (i.e. MaxRankIncrease, MinRankIncrease, RootRank) [RFC6550].

This approach assures each node with equal opportunity to insert telemetry data, despite their different distances. Although this approach may result in a lower amount of telemetry data, it will result in a better distribution of the telemetry data across nodes

and thus a more diverse set of telemetries and a more clear/wider network image.

3. Acknowledgements

TBD!

4. IANA Considerations

4.1. IETF IE Subtype INT

This document requires a number assignment in the "IEEE Std 802.15.4 IETF IE Subtype IDs" registry for IANA_IETF_IE_INT.

5. Security Considerations

Regarding the security of the INT entries, the INT protocol does not define its own security mechanisms. However, since INT fields are carried as Payload IEs, they can be encrypted and authenticated through link-layer security through CCM* with the same level of security as any other Payload IE.

INT mechanism makes use of Payload IEs in order to transfer/collect the telemetry information from network nodes. However, a malicious agent can exploit the contents of the INT Sub-IEs in order to implement a Covert Channel attack and transfer information for other purposes. Based on the INT Sub-IE Control Fields and INT request vector (Bitmap), a validation process can be applied at border router to detect and prevent possible covert/hidden channels.

Since the content of the INT sub-IE is modified at each hop, INT mechanism does not guarantee the preservation of the original telemetry information, thus creates an opportunity for a modification attack.

6. References

6.1. 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>>.

[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>>.

6.2. Informative References

- [I-D.ietf-6tisch-architecture]
Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-28 (work in progress), October 2019.
- [I-D.ietf-6tisch-coap]
Sudhaakar, R. and P. Zand, "6TiSCH Resource Management and Interaction using CoAP", draft-ietf-6tisch-coap-03 (work in progress), March 2015.
- [I-D.ietf-core-comi]
Veillette, M., Stok, P., Pelov, A., Bierman, A., and I. Petrov, "CoAP Management Interface", draft-ietf-core-comi-08 (work in progress), September 2019.
- [I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., remy@barefootnetworks.com, r., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-08 (work in progress), October 2019.
- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <<https://www.rfc-editor.org/info/rfc6550>>.
- [RFC7554] Watteyne, T., Ed., Palattella, M., and L. Grieco, "Using IEEE 802.15.4e Time-Slotted Channel Hopping (TSCH) in the Internet of Things (IoT): Problem Statement", RFC 7554, DOI 10.17487/RFC7554, May 2015, <<https://www.rfc-editor.org/info/rfc7554>>.
- [RFC8137] Kivinen, T. and P. Kinney, "IEEE 802.15.4 Information Element for the IETF", RFC 8137, DOI 10.17487/RFC8137, May 2017, <<https://www.rfc-editor.org/info/rfc8137>>.

6.3. External Informative References

- [ana2019] Alfvin, R., "802.15.4 ANA database", January 2019.

[IEEE802154]

IEEE standard for Information Technology, "IEEE Std. 802.15.4, Part. 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks", October 2019.

[IEEE802154e]

IEEE standard for Information Technology, "IEEE std. 802.15.4e, Part. 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANS) Amendment 1: MAC sublayer", April 2012.

Authors' Addresses

Abdulkadir Karaagac
Ghent University - imec
iGent Tower
Technologiepark-Zwijnaarde 126
Gent B-9052
Belgium

Email: abdulkadir.karaagac@gmail.com

Jeroen Hoebeke
Ghent University - imec
iGent Tower
Technologiepark-Zwijnaarde 126
Gent B-9052
Belgium

Email: jeroen.hoebeke@ugent.be