OPSAWG
Internet-Draft
Intended status: Standards Track
Expires: 4 September 2025

B. Claise
J. Quilbeuf
Huawei
D. Lopez
I. Dominguez
Telefonica I+D
T. Graf
Swisscom

3 March 2025

A Data Manifest for Contextualized Telemetry Data
draft-ietf-opsawg-collected-data-manifest-06

Abstract

Network platforms use Model-driven Telemetry, such as YANG-Push, to continuously stream information, including both counters and state information. This document describes the metadata that ensure that the collected data can be interpreted correctly. This document specifies the data manifest, composed of two YANG data models (the platform manifest and the data collection manifest). These YANG modules are specified at the network ~~(e.g. controller)~~ level (e.g., network controllers) to provide a model that encompasses several network platforms. The data manifest must be streamed and stored along with the data, up to the collection and analytics systems in order to keep the collected data fully exploitable by the data scientists and relevant tools.

Additionally, this document ~~proposes~~ specifies an augmentation of the YANG-Push model to include the actual collection period, in case it differs from the configured collection period.

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 4 September 2025.

Copyright Notice

Table of Contents

1. Introduction

Network platforms use Model-driven Telemetry (MDT), such as YANG-Push [RFC8641], to continuously stream information, including both counters and state information.

This document specifies what needs to be kept as metadata to ensure that the collected data can still be interpreted correctly throughout the collection and network analytics toolchain. When streaming YANG-structured data with YANG-Push ~~[RFC8641]~~, there is a semantic definition in the corresponding YANG module definition. This is the semantic information for the collected data nodes: While this

semantic is absolutely required to correctly decode and interpret the data, understanding the network platform and collection environment contexts information is equally important to interpret the data.

One part of this information is the actual collection period, as opposed to the configured collection period. On some platforms, that period can be adjusted automatically by the platform, for instance to reduce the load incurred by sending the telemetry. To later exploit the collected data, getting this actual collection period is crucial. This document defines a YANG model augmenting the YANG-Push model [RFC8641] to expose the actual collection period in Section 4~~.~~.

This document introduces the data manifest, which is composed of two YANG ~~data models~~modules, namely, the platform manifest and the data collection manifest, in order to keep the collected data exploitable by the data scientists and relevant tools.

The platform manifest contains information characterizing the platform streaming the telemetry information, while the data collection manifest contains the required information to characterize how and when the telemetry information was metered. The platform manifest is specified in Section 5. The data collection manifest is specified in Section 6.

These two ~~proposed~~ YANG modules ~~for the data manifest~~ do not expose any
new information but rather define what should be exposed by a platform streaming or storing telemetry data. Some related YANG modules have been specified to retrieve the platform capabilities such as:

* "YANG Library" [RFC8525].

* "YANG Modules Describing Capabilities for Systems and Datastore Update Notifications" [RFC9196] for the platform capabilities regarding the production and export of telemetry data.

* [I-D.claise-netconf-metadata-for-collection], which is based on [RFC9196] to define the optimal settings to stream specific items (i.e., per path).

These related YANG modules are important to discover the capabilities before applying the telemetry configuration (such as on-change subscription). Some of their content is part of the context for the streamed data.

This document covers only metadata about the collection context for the telemetry. The collected data is likely to be transformed into usable indicators for the network. The list of such transformation operations applied to the data is often called data lineage. Supplying the data lineage for the computed indicators is out of scope of this document.

To retrieve the context in which a particular piece of data was collected, three elements are necessary: the time of data emission, the originating platform and the subscription through which the data arrived. The approach ~~proposed~~ described in this document delegates the time

retrieval to the database storing the collected telemetry and focusing on providing a way to match a platform and a subscription identifier to the collection context. This is consistent with most of the YANG modules for devices, which focus on describing the current state of the device, rather than the evolution of that state through time.

2. Terminology

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.

Platform: equipment of the network able to produce telemetry.

Data manifest: The necessary data required to interpret a telemetry information.

Platform manifest: part of the data manifest that completely characterizes the platform producing the telemetry information.

Data collection manifest: part of the data manifest that completely characterizes how and when the telemetry information was metered.

Datapoint: an instance of data collected via telemetry at a specific time.

Collector: software that receives the stream of telemetry.

3. Use Cases

3.1. Network Analytics

Streamed information from network platforms is used for network analytics, incident detection, and in the closed control loop for network automation. See [I-D.ietf-nmop-terminology] for definition of some of these terms. This streamed data can be stored in a database (sometimes called a big data lake) for further analysis.

As an example, a database could store a time series representing the evolution of a specific counter collected from a network platform. When analyzing the data, a network operator/data scientist must understand the context information for these data:

* This counter definition, typically as defined in the YANG model.

* The network platform vendor, model, and OS.

* The collection parameters.

Characterizing the source used for producing the data (vendor, platform, and OS) is useful to complement the data. As an example, knowing the exact data source software specification might reveal a particularity in the observed data, explained by a specific bug, a specific bug fix, or simply a particular specific behavior. This is also necessary to ensure the reliability of the collected data. On

top of that, in particular for YANG-Push [RFC8641], it is crucial to know the set of YANG modules supported by the platform, along with their deviations. In some cases, there might even be some backwards incompatible changes in native modules (i.e., vendor proprietary modules) between one OS version to the next one. This information is captured by means of the platform manifest Section 5.

From a collection parameters point of view, the data scientists analyzing the collected data must know whether the counter was requested from the network platform as on-change or at specific cadence [RFC8641]. Indeed, an on-change collection explains why there is a single value as opposed to a time series. In case of periodic collection, this exact cadence might not be observable in the time series. Indeed, this time series might report some values as 0 or might even omit some values. The reason for this behavior might be diverse: the network platform ~~was~~ may be under stress, with a too small observation period, compared to the minimum-observed-period [I-D.claise-netconf-metadata-for-collection]. Knowing the conditions under which ~~the~~ a counter was collected and streamed (along with the platform details) ~~help~~ helps drawing the ~~right~~ informed conclusions. As an example, taking into account the value of 0 might lead to a wrong conclusion that the counter dropped to zero.

Commenté [MB1]: This expired since 2022. I would remove the citation

Commenté [MB2]: As this refers to «knowing..»

Commenté [MB3]: Some more context is needed to digest the example

3.2. New Device Onboarding

When a new device is onboarded, operators have to check that the new device streams data with YANG-Push, that the telemetry data is the right one, that the data is correctly ingested in the collection system, and finally that the data can be analyzed (compared with other similar devices). For the last point, the data manifest, which must be linked to the data up to the collection and analytics system, contains the relevant information.

Commenté [MB4]: Not sure «have to» apply for this example as yang-push is not that widely enabled.

3.3. Data Mesh Principles in Networking

The concept behind the data mesh [DataMesh] are:

* Domain Ownership: Architecturally and organizationally align business, technology, and analytical data, following the line of responsibility. The Data Mesh principles adopt the boundary of bounded context to individual data products where each domain is responsible for (and owns) its data and models.

* Data as a Product: The "Domain" owners are responsible to provide the data in useful way (discoverable through a catalog, addressable with a permanent and unique address, understandable with well-defined semantics, trustworthy and truthful, self-describing for easy consumption, interoperable by supporting standards, secure, self-contained, etc.) and should treat consumers of that data as customers. It requires and relies on the "Domain Ownership" principle.

* Self-serve Data Platform: This fosters the sharing of cross-domain data in order to create extra value.

* Federated Computational Governance: Describes the operating model

and approach to establishing global policies across a mesh of data products.

The most relevant concept for this document is the "Data as a Product" principle. The data manifest fulfills this principle as the two YANG data models, platform manifest and the data collection manifest, along with the data, provide all the necessary information in a self-describing way for easy consumption.

4. The "ietf-yp-current-period" YANG module

~~As explained earlier, the collection period is crucial information for a posteriori interpretation of the collected telemetry.~~ Some platforms will adjust the collection period depending on their capabilities and current load. The YANG module ~~proposed~~ in this section augments the "ietf-subscribed-notification" module to provide the "current-period" leaf. The value of this leaf indicates the current collection which might be different from the configured collection period.

Figure 1 contains the YANG tree diagram [RFC8340] of the "ietf-yp-current-period" module.

```
module: ietf-yp-current-period

  augment /sn:subscriptions/sn:subscription:
    +--rw current-period?   yp:centiseconds
```

Figure 1: YANG tree diagram for "ietf-yp-current-period" module

The code of the "ietf-yp-current" YANG module is given below.

```
<CODE BEGINS> file "ietf-yp-current-period@2025-02-21.yang"
module ietf-yp-current-period {
  yang-version 1.1;

  namespace "urn:ietf:params:xml:ns:yang:ietf-yp-current-period";
  prefix yp-cp;

  import ietf-subscribed-notifications {
    prefix sn;
    reference
      "RFC 8639: A YANG Data Model for Subscriptions to
       Event Notifications";
  }
  import ietf-yang-push {
    prefix yp;
    // RFC Ed.: remove revision-date, needed here because last
    // version on the server is not the ratified one
    revision-date 2019-09-09;
    reference
      "RFC 8641: Subscriptions to YANG Datastores.";
  }

  organization
    "IETF OPSAWG (Operations and Management Area) Working Group";
  contact
```

Commenté [MB5]: Any need to call the revision date?

```
        "WG Web:   <https://datatracker.ietf.org/wg/opsawg/>

         WG List:  <mailto:opsawg@ietf.org>
         Author:   Benoit Claise   <mailto:benoit.claise@huawei.com>
         Author:   Jean Quilbeuf   <mailto:jean.quilbeuf@huawei.com>
         Author:   Diego R. Lopez  <diego.r.lopez@telefonica.com>
         Author:   Ignacio Dominguez
         <ignacio.dominguezmartinez@telefonica.com>
         Author:   Thomas Graf     <thomas.graf@swisscom.com>";
      description
        "This module augments ietf-subscribed-notification and
         ietf-yang-push with the current-period statistics reporting the
         actual collection period, as opposed to the configured one.

         Copyright (c) 2025 IETF Trust and the persons identified as
         authors of the code.  All rights reserved.

         Redistribution and use in source and binary forms, with or
         without modification, is permitted pursuant to, and subject
         to the license terms contained in, the Revised BSD License
         set forth in Section 4.c of the IETF Trust's Legal Provisions
         Relating to IETF Documents
         (https://trustee.ietf.org/license-info).
         This version of this YANG module is part of RFC XXXX; see the
         RFC itself for full legal notices.  ";

      revision 2025-02-21 {
        description
          "Initial revision";
        reference
          "RFC ~~XXXX~~XXXX: A Data Manifest for Contextualized Telemetry
Data";
      }

      augment "/sn:subscriptions/sn:subscription" {
        description
          "Adds current period statistics.";
        leaf current-period {
          when '../yp:periodic';
          type yp:centiseconds;
          description
            "Period during two successive data collections, in the
             current state. Might differ from the configured period
             when the platform might increase the period
             automatically when it is overloaded.";
        }
      }
    }
    <CODE ENDS>
```

5. Platform Manifest

5.1. Overview of the Model

Figure 2 contains the YANG tree diagram ~~[RFC8340]~~ of the "ietf-platform-manifest" module.

```
module: ietf-platform-manifest
  +--ro platforms
     +--ro platform* [id]
        +--ro id                   string
        +--ro name?                string
        +--ro vendor?              string
        +--ro vendor-pen?          uint32
        +--ro software-version?    string
        +--ro software-flavor?     string
        +--ro os-version?          string
        +--ro os-type?             string
        +--ro module-set* [name]
        |  +--ro name                    string
        |  +--ro module* [name]
        |  |  +--ro name          yang:yang-identifier
        |  |  +--ro revision?     revision-identifier
        |  |  +--ro namespace     inet:uri
        |  |  +--ro location*     inet:uri
        |  |  +--ro submodule* [name]
        |  |  |  +--ro name         yang:yang-identifier
        |  |  |  +--ro revision?    revision-identifier
        |  |  |  +--ro location*    inet:uri
        |  |  +--ro feature*      yang:yang-identifier
        |  |  +--ro deviation*    -> ../../module/name
        |  +--ro import-only-module* [name revision]
        |     +--ro name          yang:yang-identifier
        |     +--ro revision      union
        |     +--ro namespace     inet:uri
        |     +--ro location*     inet:uri
        |     +--ro submodule* [name]
        |        +--ro name         yang:yang-identifier
        |        +--ro revision?    revision-identifier
        |        +--ro location*    inet:uri
        +--ro schema* [name]
        |  +--ro name           string
        |  +--ro module-set*    -> ../../module-set/name
        +--ro datastore* [name]
           +--ro name       ds:datastore-ref
           +--ro schema     -> ../../schema/name
```

Figure 2: YANG tree diagram for "ietf-platform-manifest" module

The YANG module contains a list of platform manifests (in 'platforms/platform'), indexed by the identifier of the platform. That identifier should be defined by the network manager so that each platform emitting telemetry has a unique identifier. There are several ~~ongoing~~ documents about managing the inventory of the network, e.g., [I-D.ietf-ivy-network-inventory-yang]~~, [I-D.havel-nmop-digital-map] based on [RFC8345]~~. The platform identifier should be the same as the identifier used in inventories or the 'node-id' in [RFC8345]. As an example, the identifier could be the 'sysName' from the "ietf-notification module" ~~presented~~ defined in [I-D.netana-netconf-notif-envelope].

Commenté [MB6]: Focus on adopted ones.

Commenté [MB7]: Better to use a more stable ref. RFC3418, for example.

The scope of the "ietf-platform-manifest" module is the scope of the

data collection, i.e., a given network, therefore it contains a collection of platform manifests, as opposed to the device scope, which would contain a single platform manifest.

The platform manifest is characterized by a set of parameters ('name', 'software-version', 'software-flavor', 'os-version', and 'os-type') that are aligned with the YANG Catalog [I-D.clacla-netmod-model-catalog] so that the YANG Catalog could be used to retrieve the YANG modules a posteriori. The vendor of the platform can be identified via its name 'vendor' or its PEN number 'vendor-pen', as described in [RFC9371].

a mis en forme : Surlignage

The platform manifest also includes the contents of the YANG Library [RFC8525]. That module set is particularly useful to retrieve the YANG modules associated to a subscription by analyzing the xpath filters or the subtree filters. Xpath filters are based on module names (see [RFC8639], description of leaf 'stream-xpath-filter', page 45). Subtree filters are based on namespaces.

The platform manifest is obtained by specifying the new fields defined above and mounting the YANG library module, along with the YANG ~~Revisions~~ revisions augmentations. Thus, the YANG Library part is not repeated in the YANG module for the platform manifest.

5.2. ~~YANG module~~ "ietf-platform-manifest" YANG module

This section defines the ietf-platform manifest YANG module.

```
<CODE BEGINS> file "ietf-platform-manifest@2025-02-21.yang"
module ietf-platform-manifest {
  yang-version 1.1;
  namespace "urn:ietf:params:xml:ns:yang:ietf-platform-manifest";
  prefix p-mf;

  import ietf-yang-library {
    prefix yanglib;
    reference
      "RFC8525: YANG Library";
  }

  organization
    "IETF OPSAWG (Operations and Management Area) Working Group";
  contact
    "WG Web:   <https://datatracker.ietf.org/wg/opsawg/>
     WG List:  <mailto:opsawg@ietf.org>
     Author:   Benoit Claise   <mailto:benoit.claise@huawei.com>
     Author:   Jean Quilbeuf   <mailto:jean.quilbeuf@huawei.com>
     Author:   Diego R. Lopez  <diego.r.lopez@telefonica.com>
     Author:   Ignacio Dominguez
                        <ignacio.dominguezmartinez@telefonica.com>
     Author:   Thomas Graf     <thomas.graf@swisscom.com>";
  description
    "This module describes the platform information to be used as
     context of data collection from a given network element. The
     contents of this model must be streamed along with the data
     streamed from the network element so that the platform context
     of the data collection can be retrieved later.
```

```
    The data content of this model should not change except on
    upgrade or patching of the device.

    Copyright (c) 2022—2025 IETF Trust and the persons identified as
    authors of the code.  All rights reserved.

    Redistribution and use in source and binary forms, with or
    without modification, is permitted pursuant to, and subject
    to the license terms contained in, the Revised BSD License
    set forth in Section 4.c of the IETF Trust's Legal Provisions
    Relating to IETF Documents
    (https://trustee.ietf.org/license-info).
    This version of this YANG module is part of RFC XXXX; see the
    RFC itself for full legal notices.  ";

  revision 2025-02-21 {
    description
      "Initial revision";
    reference
      "RFC xxxx: A Data Manifest for Contextualized Telemetry Data";
  }

  grouping platform-details {
    description
      "This grouping contains the information about a particular
       platform, as stored in the YANG catalog.";
    leaf name {
      type string {
        length "1..1023";
      }
      description
        "Model of the platform from which data is collected.";
    }
    leaf vendor {
      type string {
        length "1..1023";
      }
      description
        "Organization that implements that platform.";
    }
    leaf vendor-pen {
      type uint32;
      description
        "Vendor's registered Private Enterprise Number";
      reference
        "RFC9371: Registration Procedures for Private Enterprise
         Numbers (PENs)";
    }
    leaf software-version {
      type string {
        length "1..1023";
      }
      description
        "Name of the version of software. With respect to most
         network device appliances, this will be the operating system
         version. But for other YANG module implementation, this
         would be a version of appliance software. Ultimately, this
```

```
        should correspond to a version string that will be
        recognizable by the consumers of the platform.";
    }
    leaf software-flavor {
      type string {
        length "1..1023";
      }
      description
        "A variation of a specific version where YANG model support
         may be different.  Depending on the vendor, this could be a
         license, additional software component, or a feature set.";
    }
    leaf os-version {
      type string {
        length "1..1023";
      }
      description
        "Version of the operating system using this module. This is
         primarily useful if the software implementing the module is
         an application that requires a specific operating system
         version.";
    }
    leaf os-type {
      type string {
        length "1..1023";
      }
      description
        "Type of the operating system using this module. This is
         primarily useful if the software implementing the module is
         an application that requires a specific operating system
         type.";
    }
  }

  container platforms {
    config false;
    description
      "Top container including all platforms in scope. If this model
       is hosted on a single device, it should contain a single entry
       in the list. At the network level, it should contain an entry
       for every monitored platform.";
    list platform {
      key "id";
      description
        "Contains information about the platform that allows
         identifying and understanding the individual data collection
         information.";
      leaf id {
        type string {
          length "1..1023";
        }
        description
          "Identifies a given platform on the network, for instance
           the 'sysName' of the platform. The 'id' has to be unique
           within the network scope at every point in time. The same
           id can point to different platform if they are not
           simultaneously part of the network, e.g., when a device
           associated to a particular id is replaced.";
```

```
      }
      uses platform-details;
      uses yanglib:yang-library-parameters;
    }
  }
}
<CODE ENDS>
```

6. Data Collection Manifest

6.1. Overview of the Model

Commenté [MB8]: You may indicate that this part is not normative.

Figure 3 contains the YANG tree diagram [RFC8340] of the "example-collection-manifest" module. The module relies upon the YANG Schema mount [RFC8528] to reuse existing YANG modules describing the current data collection status. This module is an example as YANG Schema mount does not support design-time schema mount. Appendix C explains how the YANG tree is obtained.

```
module: example-collection-manifest
  +--ro data-collections
     +--mp data-collection* [platform-id]
        +--ro platform-id    -> /p-mf:platforms/p-mf:platform/p-mf:id
        +--ro streams/
        |  +--ro stream* [name]
        |     +--ro name           string
        |     +--ro description?   string
        +--ro filters/
        |  +--ro stream-filter* [name]
        |  |  +--ro name             string
        |  |  +--ro (filter-spec)?
        |  |     +--:(stream-subtree-filter)
        |  |     +--:(stream-xpath-filter)
        |  |        +--ro stream-xpath-filter?   yang:xpath1.0
        |  |                {xpath}?
        |  +--ro selection-filter* [filter-id]
        |     +--ro filter-id        string
        |     +--ro (filter-spec)?
        |        +--:(datastore-subtree-filter)
        |        +--:(datastore-xpath-filter)
        |           +--ro datastore-xpath-filter?   yang:xpath1.0
        |                   {sn:xpath}?
        +--ro subscriptions/
          +--ro subscription* [id]
             +--ro id                  subscription-id
             +--ro (target)
             |  +--:(stream)
             |  |  +--ro (stream-filter)?
             |  |  |  +--:(by-reference)
             |  |  |  |  +--ro stream-filter-name
             |  |  |  |          stream-filter-ref
             |  |  |  +--:(within-subscription)
             |  |  |     +--ro (filter-spec)?
             |  |  |        +--:(stream-subtree-filter)
             |  |  |        +--:(stream-xpath-filter)
             |  |  |           +--ro stream-xpath-filter?
             |  |  |                   yang:xpath1.0 {xpath}?
```

```
|  |  +--ro stream             stream-ref
|  +--:(datastore)
|     +--ro datastore             identityref
|     +--ro (selection-filter)?
|        +--:(by-reference)
|        |  +--ro selection-filter-ref
|        |          selection-filter-ref
|        +--:(within-subscription)
|           +--ro (filter-spec)?
|              +--:(datastore-subtree-filter)
|              +--:(datastore-xpath-filter)
|                 +--ro datastore-xpath-filter?
|                         yang:xpath1.0 {sn:xpath}?
+--ro stop-time?          yang:date-and-time
+--ro encoding?           encoding
+--ro receivers
|  +--ro receiver* [name]
|     +--ro name                      string
|     +--ro sent-event-records?
|     |       yang:zero-based-counter64
|     +--ro excluded-event-records?
|     |       yang:zero-based-counter64
|     +--ro state                     enumeration
+--ro (update-trigger)?
|  +--:(periodic)
|  |  +--ro periodic!
|  |     +--ro period         centiseconds
|  |     +--ro anchor-time?   yang:date-and-time
|  +--:(on-change) {on-change}?
|     +--ro on-change!
|        +--ro dampening-period?   centiseconds
|        +--ro sync-on-start?      boolean
|        +--ro excluded-change*    change-type
+--ro current-period?     yp:centiseconds
```

Figure 3: YANG tree diagram for example-collection-manifest module

The 'data-collections' container contains the information related to each YANG-Push subscription. As for the platform manifest, these subscriptions are indexed by the 'platform-id', so that all subscriptions in the network can be represented at the network level without any conflict.

As most of the information related to YANG-push subscription [RFC8639] and [RFC8641] is stored in the "ietf-yang-push" module, these modules are mounted. These modules have a part common to all subscriptions of the platform, stored in the 'streams' and 'filters' containers. The information about subscriptions themselves are stored in the 'subscriptions/subscription' list, indexed by a subscription identifier.

In the subscription object, the 'current-period' indicates the period currently used between two updates. That leaf can only be present when the subscription is periodic. The current period might differ from the requested period if the platform implements a mechanism to increase the collection period when it is overloaded. Having the current period information is crucial to understand if telemetry is missing because of a bug or a packet loss or simply because it was

dynamically adjusted by the platform.

The 'current-period' data node is added by the module 'ietf-data-collection-manifest-statistics' presented in Section 4. This example module augments the subscription list from the module 'ietf-subscribed-notifications'. It is mounted as well via the YANG Schema Mount mechanism. The module for the data collection manifest is presented in Section 6.2.

6.2. The "example-collection-manifest" YANG module

This section includes the code of the "example-collection-manifest" YANG module. Additionally, it defines the extension data file for YANG schema mount. The data collection manifest MUST conform to the model obtained by combining these two specifications.

Commenté [MB9]: Weird use of normative language as this is about an example module

```
module example-collection-manifest {
  yang-version 1.1;
  namespace "http://example.com/example-data-collection-manifest";
  prefix ex-d-mf;

  import ietf-platform-manifest {
    prefix p-mf;
    reference
      "RFC XXXX: A Data Manifest for Contextualized Telemetry Data
Title to be completed";
  }
  import ietf-yang-schema-mount {
    prefix yangmnt;
    reference
      "RFC8528: YANG Schema Mount";
  }

  organization
    "IETF OPSAWG (Operations and Management Area) Working Group";
  contact
    "WG Web:   <https://datatracker.ietf.org/wg/opsawg/>
     WG List:  <mailto:opsawg@ietf.org>

     Author:   Benoit Claise   <mailto:benoit.claise@huawei.com>
     Author:   Jean Quilbeuf   <mailto:jean.quilbeuf@huawei.com>
     Author:   Diego R. Lopez  <diego.r.lopez@telefonica.com>
     Author:   Ignacio Dominguez
                          <ignacio.dominguezmartinez@telefonica.com>
     Author:   Thomas Graf     <thomas.graf@swisscom.com>";
  description
    "This module describes the context of data collection from a
     given network element. The contents of this model must be
     streamed along with the data streamed from the network
     element so that the context of the data collection can
     be retrieved later.

     This module must be completed with
     ietf-platform-manifest
     to capture the whole context of a data collection session.

     The key words 'MUST', 'MUST NOT', 'REQUIRED', 'SHALL',
```

Commenté [MB10]: Per 8407bis guidance

```
        'SHALL NOT', 'SHOULD', 'SHOULD NOT', 'RECOMMENDED',
        'NOT RECOMMENDED', 'MAY', and 'OPTIONAL' in this document
        are to be interpreted as described in BCP 14 (RFC 2119)
        (RFC 8174) when, and only when, they appear in all
        capitals, as shown here.

        Copyright (c) 2025 IETF Trust and the persons identified as
        authors of the code.  All rights reserved.

        Redistribution and use in source and binary forms, with or
        without modification, is permitted pursuant to, and subject
        to the license terms contained in, the Revised BSD License
        set forth in Section 4.c of the IETF Trust's Legal Provisions
        Relating to IETF Documents
        (https://trustee.ietf.org/license-info).
        This version of this YANG module is part of RFC XXXX; see the
        RFC itself for full legal notices.  ";

    revision 2025-02-21 {
      description
        "Initial revision";
      reference
        "RFC XXXX: A Data Manifest for Contextualized Telemetry Data";
    }

    container data-collections {
      config false;
      description
        "Contains the configuration and statistics for the collected
         data, per node in the network.";

      list data-collection {
        key "platform-id";
        description
          "Defines the information for each collected object.";
        leaf platform-id {
          type leafref {
            path "/p-mf:platforms/p-mf:platform/p-mf:id";
          }
          description
            "Identifier of the platform collecting the data. This
             identifier is the same as the one in the platform
             manifest.";
        }
        yangmnt:mount-point "yang-push-collection" {
          description
            "This mount point MUST mount the following modules and their
             dependencies:
               * ietf-subscribed-notifications
               * ietf-yang-push
               * ietf-yp-current-period.
             This mount point MUST NOT mount any other modules.";
          reference
            "RFC8639: Subscription to YANG Notifications
             RFC8641: Subscription to YANG Notifications for datastore
                      updates";
        }
      }
```

Commenté [MB11]: Weird in an example module

Commenté [MB12]: Weird in an example

Commenté [MB13]: Weird in an example module

```
      }
    }
```

7. Data Manifest and the Collected Data

This section focuses on ~~relating~~ associating the collected data to the data manifest. As this document specifically focuses on giving context on data collected via streamed telemetry, it is assumed that a streaming telemetry system is available. Another premise of this document is the storage of the collected data into a database for later exploitation. ~~It~~ This document ~~is assumed~~ assumes that such a database exists and can be used for storing the data manifest.

7.1. Collecting the Data Manifest

The data manifest MUST be streamed and stored along with the collected data. In case the collected data are moved to a different place (typically a database), the companion data manifest MUST follow the collected data. Storing the collected data without the companion data manifest might prevent the correct interpretation of the collected data. The data manifest MUST be updated when the data manifest information changes, for example, when a router is upgraded, when a new telemetry subscription is configured, or when the telemetry subscription parameters change. The data manifest can itself be considered as a time series, and stored in a similar fashion to the collected data.

This document recommends reusing ~~the~~ existing telemetry systems (in-band approach) in order to lower the efforts for implementing this approach. To enable a platform supporting streaming telemetry to also support the data manifest, it is sufficient that this platform supports the models from Sections 5 and 6. The collection of the data manifest MUST be explicitly configured by the collector by requesting the relevant subscriptions. These subscriptions MUST include the platform manifest and the data collection manifest, possibly limited to the subscriptions for which the context needs to be retrieved a posteriori. Appendix B shows how the in-band approach would work while storing to a time-series database (TSDB).

Each type of manifest has its own rough frequency update, i.e.~~,~~ at reboot for the platform manifest and when subscriptions are modified for the data collection manifest. The data manifest SHOULD be streamed with the YANG-Push on-change feature [RFC8641] (also called event-driven telemetry) whenever possible.

a mis en forme : Surlignage

A platform manifest is likely to remain the same until the platform is updated. Thus, the platform manifest only needs to be collected once per streaming session and updated after a platform reboot. The "subscription-terminated" (Section 2.7.3 of [RFC8639]) will indicate to the collector that the platform rebooted. The collector MUST then collect the potential update of the platform manifest on re-establishment of the subscription. Using the on-change feature enables to capture dynamic changes to the platform manifest as well, if any.

Regarding the data manifest, the elements common to all subscriptions, such as the stream definitions and the common filters might be updated less frequently than the subscriptions. Relying on YANG-Push on-change feature enables keeping an up-to-date version of the data collection manifest.

The underlying time series database should accommodate the various rates at which different parts of the data manifest are updated. In particular, storing the platform manifest should be optimized to avoid duplicating repeated content and only storing a new version when there is a change in the manifest.

7.2. Mapping Collected Data to the Data Manifest

As explained in the introduction, three elements are necessary to identify the data manifest associated to a datapoint:

* the time at which the data was sent from the device,

* the originating platform sending the data, and

* the identifier of the subscription that produced the data.

~~This~~ These elements can be either known to the collector, if it is the one configuring the collection, or retrieved via dedicated headers as proposed, e.g., in [I-D.netana-netconf-notif-envelope]. In order to enable a posteriori retrieval of the data manifest associated to a datapoint, the collector MUST keep the subscription identifier and platform identifier in the metadata of the collected values.

With this information, to retrieve the data manifest from a datapoint, the following happens:

Commenté [MB14]: Which one?

* The subscription identifier, platform identifier and timestamp~~time stamp~~ of the data are retrieved from the datapoint metadata

* The platform manifest for that datapoint is obtained by looking up the latest version before the timestamp~~time stamp~~ matching the platform identifier.

* The data collection manifest for that datapoint is obtained by looking up the latest version before the time matching the platform identifier and the subscription identifier.

The reliability of the collection of the data manifest is the same as the reliability of the data collection itself, since the data manifest is like any other data.

7.3. Operational Considerations

It is expected that the data manifest is streamed directly from the network equipment, along with YANG-Push [RFC8641] data. However, if the network equipment streaming telemetry does not yet support the YANG modules from the data manifest specified in this document, the

telemetry collector could populate the data manifest from available information collected from the platform. This latter option requires efforts on the telemetry collector side, as the information gathered in the data manifest proposed in this document could be scattered among various standard and vendor-specific YANG modules [RFC8199], that depend on the platform.

8. Example

Figure 4 shows an example of both a Platform manifest and corresponding data collection manifests. The list of YANG modules in the 'yang-library' container is kept empty for brevity.

```
{
  "ietf-platform-manifest:platforms": {
    "platform": [
      {
        "id": "PE1",
        "name": "PE1",
        "vendor": "ACME",
        "vendor-pen": 32473,
        "software-version": "3.14",
        "os-version": "2.79",
        "os-type": "ACME OS"
      }
    ]
  },
  "example-collection-manifest:data-collections": {
    "data-collection": [
      {
        "platform-id": "PE1",
        "ietf-subscribed-notifications:subscriptions": {
          "subscription": [
            {
              "id": 4242,
              "ietf-yang-push:datastore":
                                    "ietf-datastores:operational",
              "ietf-yang-push:datastore-xpath-filter":
                     "/ietf-interfaces:interfaces/interface/enabled",

              "ietf-yang-push:on-change": {},
              "receivers": {
                "receiver": [
                  {
                    "name": "yp-collector",
                    "state": "active"
                  }
                ]
              }
            },
            {
              "id": 4243,
              "ietf-yang-push:datastore":
                                    "ietf-datastores:operational",
              "ietf-yang-push:datastore-xpath-filter":
        "/ietf-interfaces:interfaces/interface/statistics/in-octets",
              "ietf-yang-push:periodic": {
```

Commenté [MB15]: Please validate all the examples

```
                "period": 10000
              },
            "ietf-yp-current-period:current-period": 20000,
              "receivers": {
                "receiver": [
                  {
                    "name": "yp-collector",
                    "state": "active"
                  }
                ]
              }
            }
          ]
        }
      }
    ]
  }
}
```

Figure 4: Example of data manifest

Figure 4 contains the data collection manifest for two XPaths subscriptions. With the data collection manifest for the first one, with subscription identifier 4242, the exact semantics of the collected path, here the administrative status of the network interfaces, can be obtained by looking up the module in the yang-library of the corresponding platform manifest, in order to obtain the exact revision of ietf-interfaces used at collection time. Also, the "on-change" container indicates that data will be sent only if there is a change, thus not receiving data indicates that the administrative status of the interface did not change.

The other example of data collection manifest, with subscription identifier 4243, shows how a periodic subscription is reported. In that example, the 'current-period' indicates that the requested period of 10s (1000 centiseconds) could not be attained and is now of 20s, for instance because the device is overloaded.

Appendix D gives the command line for validating this example using [yanglint].

9. Security Considerations

Commenté [MB16]: Please use the latest sec template. See rfc8407bis

The YANG modules specified in this document define a schema for data that is designed to be accessed via network management protocols such as NETCONF [RFC6241] or RESTCONF [RFC8040]. The lowest NETCONF layer is the secure transport layer, and the mandatory-to-implement secure transport is Secure Shell (SSH) [RFC6242]. The lowest RESTCONF layer is HTTPS, and the mandatory-to-implement secure transport is TLS [RFC8446].

The Network Configuration Access Control Model (NACM) [RFC8341] provides the means to restrict access for particular NETCONF or RESTCONF users to a preconfigured subset of all available NETCONF or RESTCONF protocol operations and content.

Some of the readable data nodes in this YANG module may be considered sensitive or vulnerable in some network environments. It is thus

important to control read access (e.g., via get, get-config, or notification) to these data nodes. These are the subtrees and data nodes and their sensitivity/vulnerability:

* ietf-plaftorm-manifest:platforms/platform contains details about the platform that an attacker could use to find the known vulnerabilities of the platform.

As the present approach reuses an existing telemetry system, the security considerations lie with the new content divulged in the new manifests. Appropriate access control filters must be associated to the corresponding leafs and containers, as well as the databases storing them.

The integrity and provenance of the data of the collection manifest can be ensured by a signing mechanism such as [I-D.lopez-opsawg-yang-provenance].

10. IANA Considerations

RFC Ed.: replace XXXX with actual RFC number and remove this note.

IANA is requested to register the following URIs in the "ns" subregistry within the "IETF XML Registry" [RFC3688]:

```
URI: urn:ietf:params:xml:ns:yang:ietf-platform-manifest
Registrant Contact:  The IESG.
XML: N/A; the requested URI is an XML namespace.

URI: urn:ietf:params:xml:ns:yang:ietf-yp-current-period
Registrant Contact:  The IESG.
XML: N/A; the requested URI is an XML namespace.
```

IANA is requested to register the following YANG modules in the "YANG Module Names" subregistry [RFC6020] within the "YANG Parameters" registry.

```
Name: ietf-platform-manifest
Maintained by IANA?  N
Namespace: urn:ietf:params:xml:ns:yang:ietf-platform-manifest
Prefix: p-mf
Reference: RFC XXXX

Name: ietf-yp-current-period
Maintained by IANA?  N
Namespace: urn:ietf:params:xml:ns:yang:ietf-yp-current-period
Prefix: yp-cp
Reference: RFC XXXX
```

11. Contributors

12. Open Issues

This section is to be removed before publishing as an RFC.

* Do we want to handle the absence of values, i.e. add information about missed collection or errors in the collection context ? It

could also explain why some values are missing. On the other hand, this might also be out scope. CLOSED: the goal of the manifest is to be able to detect miscollection a posteriori. Assurance of the metric collection is out of scope and could be done via an external mechanism such as SAIN.

* Henk: how does this interact with SBOM effort? CLOSED: SBOM is another kind of manifest, we are focusing here on data collection.

* What is the link with the RFC8345 NodeId and IVY? CLOSED: added text.

* Handling of deletion in [I-D.kll-yang-label-tsdb]. CLOSED: out of scope

13. 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>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, DOI 10.17487/RFC6242, June 2011, <https://www.rfc-editor.org/info/rfc6242>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[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>.

[RFC8341] Bierman, A. and M. Bjorklund, "Network Configuration Access Control Model", STD 91, RFC 8341, DOI 10.17487/RFC8341, March 2018, <https://www.rfc-editor.org/info/rfc8341>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[RFC8525] Bierman, A., Bjorklund, M., Schoenwaelder, J., Watsen, K., and R. Wilton, "YANG Library", RFC 8525, DOI 10.17487/RFC8525, March 2019, <https://www.rfc-editor.org/info/rfc8525>.

[RFC8528] Bjorklund, M. and L. Lhotka, "YANG Schema Mount", RFC 8528, DOI 10.17487/RFC8528, March 2019, <https://www.rfc-editor.org/info/rfc8528>.

[RFC8639] Voit, E., Clemm, A., Gonzalez Prieto, A., Nilsen-Nygaard, E., and A. Tripathy, "Subscription to YANG Notifications", RFC 8639, DOI 10.17487/RFC8639, September 2019, <https://www.rfc-editor.org/info/rfc8639>.

[RFC8641] Clemm, A. and E. Voit, "Subscription to YANG Notifications for Datastore Updates", RFC 8641, DOI 10.17487/RFC8641, September 2019, <https://www.rfc-editor.org/info/rfc8641>.

14. Informative References

[DataMesh] "Datamesh Architecture", <https://www.datamesh-architecture.com/>.

[I-D.clacla-netmod-model-catalog]
Clarke, J. and B. Claise, "YANG module for yangcatalog.org", Work in Progress, Internet-Draft, draft-clacla-netmod-model-catalog-03, 3 April 2018, <https://datatracker.ietf.org/doc/html/draft-clacla-netmod-model-catalog-03>.

[I-D.claise-netconf-metadata-for-collection]
Claise, B., Nayyar, M., and A. R. Sesani, "Per-Node Capabilities for Optimum Operational Data Collection", Work in Progress, Internet-Draft, draft-claise-netconf-metadata-for-collection-03, 25 January 2022, <https://datatracker.ietf.org/doc/html/draft-claise-netconf-metadata-for-collection-03>.

[I-D.havel-nmop-digital-map]
Havel, O., Claise, B., de Dios, O. G., Elhassany, A., and T. Graf, "Modeling the Digital Map based on RFC 8345: Sharing Experience and Perspectives", Work in Progress, Internet-Draft, draft-havel-nmop-digital-map-02, 21 October 2024, <https://datatracker.ietf.org/doc/html/draft-havel-nmop-digital-map-02>.

[I-D.ietf-ivy-network-inventory-yang]
Yu, C., Belotti, S., Bouquier, J., Peruzzini, F., and P. Bedard, "A Base YANG Data Model for Network Inventory", Work in Progress, Internet-Draft, draft-ietf-ivy-network-inventory-yang-05, 28 February 2025, <https://datatracker.ietf.org/api/v1/doc/document/draft-ietf-ivy-network-inventory-yang/>.

[I-D.ietf-nmop-terminology]
Davis, N., Farrel, A., Graf, T., Wu, Q., and C. Yu, "Some Key Terms for Network Fault and Problem Management", Work in Progress, Internet-Draft, draft-ietf-nmop-terminology-12, 22 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-nmop-terminology-12>.

[I-D.kll-yang-label-tsdb]
Larsson, K., "Mapping YANG Data to Label-Set Time Series", Work in Progress, Internet-Draft, draft-kll-yang-label-tsdb-00, 18 October 2023, <https://datatracker.ietf.org/doc/html/draft-kll-yang-label-tsdb-00>.

[I-D.lopez-opsawg-yang-provenance]
Lopez, D., Pastor, A., Feng, A. H., Mendez, A., Birkholz, H., and S. Garcia, "Applying COSE Signatures for YANG Data Provenance", Work in Progress, Internet-Draft, draft-lopez-opsawg-yang-provenance-05, 26 February 2025, <https://datatracker.ietf.org/doc/html/draft-lopez-opsawg-yang-provenance-05>.

[I-D.netana-netconf-notif-envelope]
Feng, A. H., Francois, P., Graf, T., and B. Claise, "Extensible YANG Model for YANG-Push Notifications", Work in Progress, Internet-Draft, draft-netana-netconf-notif-

envelope-02, 28 January 2025,
<https://datatracker.ietf.org/doc/html/draft-netana-netconf-notif-envelope-02>.

[RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688, DOI 10.17487/RFC3688, January 2004, <https://www.rfc-editor.org/info/rfc3688>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <https://www.rfc-editor.org/info/rfc6020>.

[RFC8199] Bogdanovic, D., Claise, B., and C. Moberg, "YANG Module Classification", RFC 8199, DOI 10.17487/RFC8199, July 2017, <https://www.rfc-editor.org/info/rfc8199>.

[RFC8340] Bjorklund, M. and L. Berger, Ed., "YANG Tree Diagrams", BCP 215, RFC 8340, DOI 10.17487/RFC8340, March 2018, <https://www.rfc-editor.org/info/rfc8340>.

[RFC8343] Bjorklund, M., "A YANG Data Model for Interface Management", RFC 8343, DOI 10.17487/RFC8343, March 2018, <https://www.rfc-editor.org/info/rfc8343>.

[RFC8345] Clemm, A., Medved, J., Varga, R., Bahadur, N., Ananthakrishnan, H., and X. Liu, "A YANG Data Model for Network Topologies", RFC 8345, DOI 10.17487/RFC8345, March 2018, <https://www.rfc-editor.org/info/rfc8345>.

[RFC9196] Lengyel, B., Clemm, A., and B. Claise, "YANG Modules Describing Capabilities for Systems and Datastore Update Notifications", RFC 9196, DOI 10.17487/RFC9196, February 2022, <https://www.rfc-editor.org/info/rfc9196>.

[RFC9371] Baber, A. and P. Hoffman, "Registration Procedures for Private Enterprise Numbers (PENs)", RFC 9371, DOI 10.17487/RFC9371, March 2023, <https://www.rfc-editor.org/info/rfc9371>.

[yanglint] "Yanglint", <https://github.com/CESNET/libyang>.