

NETCONF  
Internet-Draft  
Intended status: Standards Track  
Expires: April 21, 2019

T. Zhou  
G. Zheng  
Huawei  
E. Voit  
Cisco Systems  
A. Clemm  
Huawei  
A. Bierman  
YumaWorks  
October 18, 2018

**Subscription to Multiple Stream Originators**  
**draft-zhou-netconf-multi-stream-originators-03**

Abstract

This document describes the distributed data collection mechanism that allows multiple data streams to be managed using a single subscription. Specifically, multiple data streams are pushed directly to the collector without passing through a broker for internal consolidation.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 21, 2019.

## Copyright Notice

Copyright (c) 2018 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

|                             |   |                    |
|-----------------------------|---|--------------------|
| <a href="#">1.</a>          | Introduction . . . . .  | <a href="#">2</a>  |
| <a href="#">2.</a>          | Use Cases . . . . .   | <a href="#">3</a>  |
| 2.1.                        | Use Case 1: Data Collection from Devices with Main-board and Line-cards . . . . . | <a href="#">3</a>  |
| <a href="#">2.2.</a>        | Use Case 2: IoT Data Collection . . . . .   | <a href="#">4</a>  |
| <a href="#">3.</a>          | Terminologies . . . . .   | <a href="#">5</a>  |
| <a href="#">4.</a>          | Solution Overview . . . . .   | <a href="#">6</a>  |
| <a href="#">5.</a>          | Subscription Decomposition . . . . .  | <a href="#">8</a>  |
| <a href="#">6.</a>          | Publication Composition . . . . .   | <a href="#">9</a>  |
| <a href="#">7.</a>          | Subscription State Change Notifications . . . . .                                 | <a href="#">10</a> |
| <a href="#">8.</a>          | IANA Considerations . . . . .   | <a href="#">10</a> |
| <a href="#">9.</a>          | Security Considerations . . . . .   | <a href="#">10</a> |
| <a href="#">10.</a>         | Acknowledgements . . . . .  | <a href="#">10</a> |
| <a href="#">11.</a>         | References . . . . .  | <a href="#">10</a> |
| <a href="#">11.1.</a>       | Normative References . . . . .  | <a href="#">11</a> |
| <a href="#">11.2.</a>       | Informative References . . . . .  | <a href="#">11</a> |
| <a href="#">Appendix A.</a> | Change Log . . . . .  | <a href="#">12</a> |
|                             | Authors' Addresses . . . . .  | <a href="#">12</a> |

## [1.](#) Introduction

Streaming telemetry refers to sending a continuous stream of operational data from a device to a remote receiver. This provides an ability to monitor a network from remote and to provide network analytics. Devices generate telemetry data and push that data to a collector for further analysis. By streaming the data, much better performance, finer-grained sampling, monitoring accuracy, and bandwidth utilization can be achieved than with polling-based alternatives.



YANG-Push [[I-D.ietf-netconf-yang-push](#)] defines a transport-independent subscription mechanism for datastore updates, in which a subscriber can subscribe to a stream of datastore updates from a server, or update provider. The current design involves subscription to a single push server. This conceptually centralized model encounters efficiency limitations in cases where the data sources are themselves distributed, such as line cards in a piece of network equipment. In such cases, it will be a lot more efficient to have each data source (e.g., each line card) originate its own stream of updates, rather than requiring updates to be tunneled through a central server where they are combined. What is needed is a distributed mechanism that allows to directly push multiple individual data substreams, without needing to first pass them through an additional processing stage for internal consolidation, but still allowing those substreams to be managed and controlled via a single subscription.

This document will describe such distributed data collection mechanism and how it can work by extending existing YANG-Push mechanism. The proposal is general enough to fit many scenarios.

## **[2.](#) Use Cases**

### **[2.1.](#) Use Case 1: Data Collection from Devices with Main-board and Line-cards**

For data collection from devices with main-board and line-cards, existing YANG-Push solutions consider only one push server typically reside in the main board. As shown in the following figure, data are collected from line cards and aggregate to the main board as one consolidated stream. So the main board can easily become the performance bottle-neck. The optimization is to apply the distributed data collection mechanism which can directly push data from line cards to a collector. On one hand, this will reduce the cost of scarce compute and memory resources on the main board for data processing and assembling. On the other hand, distributed data push can off-load the streaming traffic to multiple interfaces.



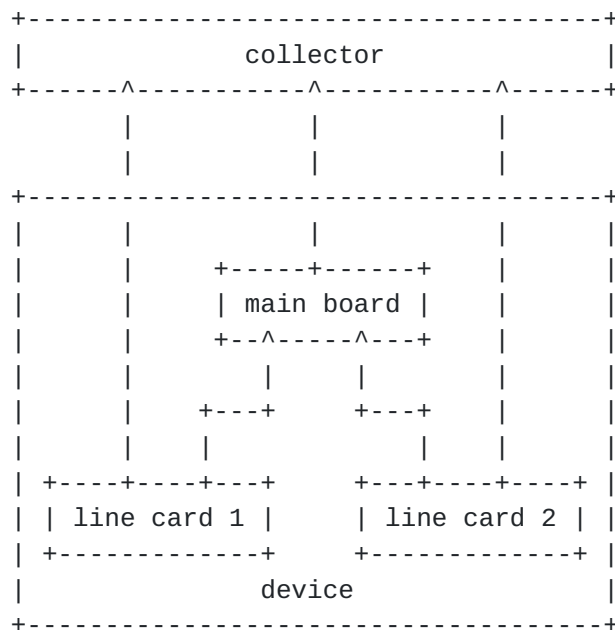


Fig. 1 Data Collection from Devices with Main-board and Line-cards

## 2.2. Use Case 2: IoT Data Collection

In the IoT data collection scenario, as shown in the following figure, collector usually cannot access to IoT nodes directly, but is isolated by the border router. So the collector subscribes data from the border router, and let the border router to disassemble the subscription to corresponding IoT nodes. The border router is typically the traffic convergence point. It's intuitive to treat the border router as a broker assembling the data collected from the IoT nodes and forwarding to the collector[I-D.ietf-core-coap-pubsub]. However, the border router is not so powerful on data assembling as a network device. It's more efficient for the collector, which may be a server or even a cluster, to assemble the subscribed data if possible. In this case, push servers that reside in IoT nodes can stream data to the collector directly while traffic only passes through the border router.



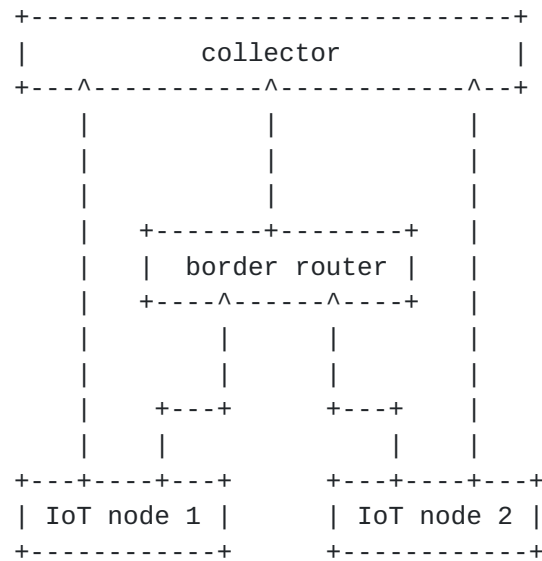


Fig. 2 IoT Data Collection

### 3. Terminologies

**Subscriber:** generates the subscription instructions to express what and how the collector want to receive the data

**Receiver:** is the target for the data publication.

**Publisher:** pushes data to the receiver according to the subscription information.

**Subscription Server:** which manages capabilities that it can provide to the subscriber.

**Global Subscription:** the subscription requested by the subscriber. It may be decomposed into multiple Component Subscriptions.

**Component Subscription:** is the subscription that defines the data from each individual telemetry source which is managed and controlled by a single Subscription Server.

**Global Capability:** is the overall subscription capability that the group of Publishers can expose to the Subscriber.

**Component Capability:** is the subscription capability that each Publisher can expose to the Subscriber.

**Master Publication Channel:** the session between the Master Publisher and the Receiver.





Agent Publication Channel: the session between the Agent Publisher and the Receiver.

#### **4. Solution Overview**

All the use cases described in the previous section are very similar on the data subscription and publication mode, hence can be abstracted to the following generic distributed data collection framework, as shown in the following figure.

A Collector usually includes two components,

- o the Subscriber generates the subscription instructions to express what and how the collector want to receive the data;
- o the Receiver is the target for the data publication.

For one subscription, there may be one to many receivers. And the subscriber does not necessarily share the same address with the receivers.

In this framework, the Publisher pushes data to the receiver according to the subscription information. The Publisher has the Master role and the Agent role. Both the Master and the Agent include the Subscription Server which actually manages capabilities that it can provide to the subscriber.

The Master knows all the capabilities that the attached Agents and itself can provide, and exposes the Global Capability to the Collector. The Collector cannot see the Agents directly, so it will only send the Global Subscription information to the Master. The Master disassembles the Global Subscription to multiple Component Subscriptions, each involving data from a separate telemetry source. The Component Subscriptions are then distributed to the corresponding Agents.

When data streaming, the Publisher collects and encapsulates the packets per the Component Subscription, and pushes the piece of data which can serve directly to the designated data Collector. The Collector is able to assemble many pieces of data associated with one Global Subscription, and can also deduce the missing pieces of data.



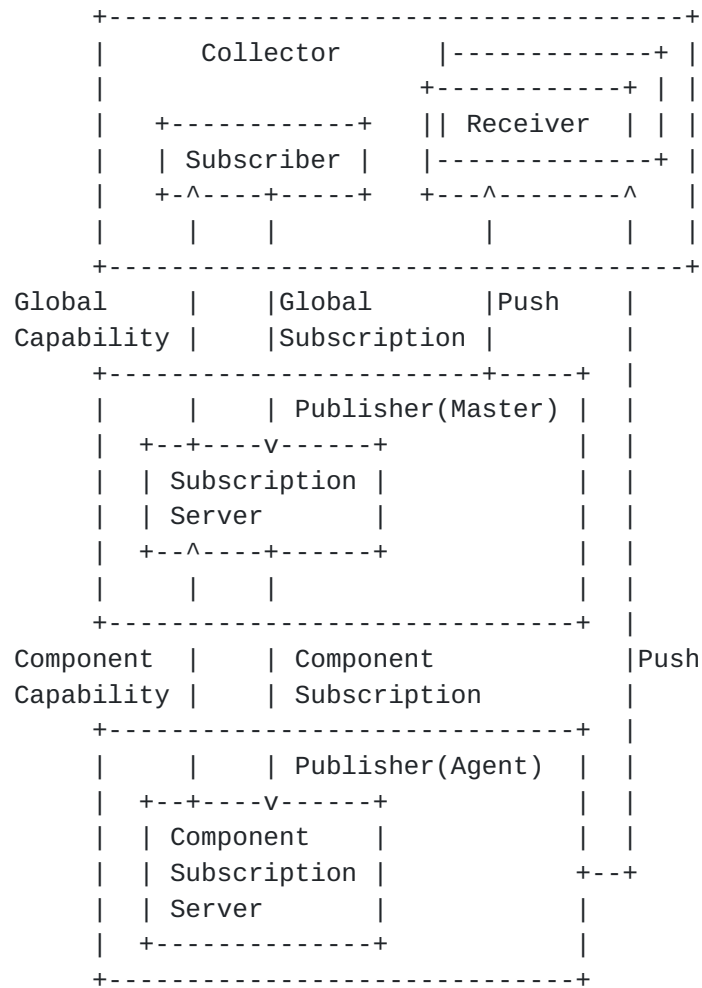


Fig. 3 The Generic Distributed Data Collection Framework

Master and Agents may interact with each other in several ways:

- o Agents need to have a registration or announcement handshake with the Master, so the Master is aware of them and of life-cycle events (such as Agent appearing and disappearing).
- o Contracts are needed between the Master and each Agent on the Component Capability, and the format for streaming data structure.
- o The Master relays the component subscriptions to the Agents.
- o The Agents indicate status of Component Subscriptions to the Master. The status of the overall subscription is maintained by the Master. The Master is also responsible for notifying the subscriber in case of any problems of Component Subscriptions.



Any technical mechanisms or protocols used for the coordination of operational information between Master and Agent is out-of-scope of the solution. We will need to instrument the results of this coordination on the Master Node.

## **5. Subscription Decomposition**

Since Agents are invisible to the Collector, the Collector can only subscribe to the Master. This requires the Master to:

1. expose the Global Capability that can be served by multiple Publishers;
2. disassemble the Global Subscription to multiple Component Subscriptions, and distribute them to the corresponding telemetry sources;
3. notify on changes when portions of a subscription moving between different Agents over time.

To achieve the above requirements, the Master need a Global Capability description which is typically the YANG [[RFC7950](#)] data model. This global YANG model is provided as the contract between the Master and the Collector. Each Agent associating with the Master owns a local YANG model to describe the Component Capabilities which it can serve as part of the Global Capability. All the Agents need to know the namespace associated with the Master.

The Master also need a data structure, typically a Resource-Location Table, to keep track of the mapping between the resource and the corresponding location of the Subscription Server which commits to serve the data. When a Global Subscription request arrives, the Master will firstly extract the filter information from the request. Consequently, according to the Resource-Location Table, the Global Subscription can be disassembled into multiple Component Subscriptions, and the corresponding location can be associated.

The decision whether to decompose a Global Subscription into multiple Component Subscriptions rests with the Resource-Location Table. A Master can decide to not decompose a Global Subscription at all and push a single stream to the receiver, because the location information indicates the Global Subscription can be served locally by the Master. Similarly, it can decide to entirely decompose a Global Subscription into multiple Component Subscriptions that each push their own streams, but not from the Master. It can also decide to decompose the Global Subscription into several Component Subscriptions and retain some aspects of the Global Subscription itself, also pushing its own stream.



Component Subscriptions belonging to the same Global Subscription MUST NOT overlap. The combination of all Component Subscriptions MUST cover the same range of nodes as the Global Subscription. Also, the same subscription settings apply to each Component Subscription, i.e., the same receivers, the same time periods, the same encodings are applied to each Component Subscription per the settings of the Global Subscription.

Each Component Subscription in effect constitutes a full-fledged subscription, with the following constraints:

- o Component subscriptions are system-controlled, i.e. managed by the Master, not by the subscriber.
- o Component subscription settings such as time periods, dampening periods, encodings, receivers adopt the settings of their Global Subscription.
- o The life-cycle of the Component Subscription is tied to the life-cycle of the Global Subscription. Specifically, terminating/removing the Global Subscription results in termination/removal of Component Subscriptions.
- o The Component Subscriptions share the same Subscription ID as the Global Subscription.

## **6. Publication Composition**

The Publisher collects data and encapsulates the packets per the Component Subscription. There are several potential encodings, including XML, JSON, CBOR and GPB. The format and structure of the data records are defined by the YANG schema, so that the composition at the Receiver can benefit from the structured and hierarchical data instance.

The Receiver is able to assemble many pieces of data associated with one subscription, and can also deduce the missing pieces of data. The Receiver recognizes data records associated with one subscription according the Subscription ID. Data records generated per one subscription are assigned with the same Subscription ID.

For the time series data stream, records are produced periodically from each stream originator. The message arrival time varies because of the distributed nature of the publication. The Receiver assembles data generated at the same time period based on the recording time consisted in each data record. In this case, time synchronization is required for all the Publishers.





To check the integrity of the data generated from different Publishers at the same time period, the Message Generator ID [[I-D.ietf-netconf-notification-messages](#)] is helpful. This requires the Subscriber to know the number of Component Subscriptions which the Global Subscription is decomposed to. For the dynamic subscription, the response of the "establish-subscription" and "modify-subscription" RPC defined in [[I-D.ietf-netconf-subscribed-notifications](#)] can include a list of Message Generator IDs to indicate how the Global Subscription is decomposed into several Component Subscriptions. The "subscription-started" and "subscription-modified" notification defined in [[I-D.ietf-netconf-subscribed-notifications](#)] can also include a list of Message Generator IDs to notify the current Publishers for the corresponding Global Subscription.

## **[7.](#) Subscription State Change Notifications**

In addition to sending event records to receivers, the Master MUST also send subscription state change notifications [[I-D.ietf-netconf-subscribed-notifications](#)] when events related to subscription management have occurred. All the subscription state change notifications MUST be delivered by the Master Publication Channel which is the session between the Master Publisher and the Receiver.

When the subscription decomposition result changed, the "subscription-modified" notification will be sent to indicate the new a list of Publishers.

## **[8.](#) IANA Considerations**

TBD

## **[9.](#) Security Considerations**

It's expected to reuse the existing secure transport layer protocols, such as TLS [[RFC5246](#)] and DTLS [[RFC6347](#)], to secure the telemetry stream.

## **[10.](#) Acknowledgements**

TBD

## **[11.](#) References**



### **11.1. 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>>.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", [RFC 5246](#), DOI 10.17487/RFC5246, August 2008, <<https://www.rfc-editor.org/info/rfc5246>>.
- [RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", [RFC 6347](#), DOI 10.17487/RFC6347, January 2012, <<https://www.rfc-editor.org/info/rfc6347>>.
- [RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", [RFC 7950](#), DOI 10.17487/RFC7950, August 2016, <<https://www.rfc-editor.org/info/rfc7950>>.

### **11.2. Informative References**

- [I-D.ietf-core-coap-pubsub] Koster, M., Keranen, A., and J. Jimenez, "Publish-Subscribe Broker for the Constrained Application Protocol (CoAP)", [draft-ietf-core-coap-pubsub-05](#) (work in progress), July 2018.
- [I-D.ietf-netconf-notification-messages] Voit, E., Birkholz, H., Bierman, A., Clemm, A., and T. Jenkins, "Notification Message Headers and Bundles", [draft-ietf-netconf-notification-messages-04](#) (work in progress), August 2018.
- [I-D.ietf-netconf-subscribed-notifications] Voit, E., Clemm, A., Prieto, A., Nilsen-Nygaard, E., and A. Tripathy, "Customized Subscriptions to a Publisher's Event Streams", [draft-ietf-netconf-subscribed-notifications-17](#) (work in progress), September 2018.
- [I-D.ietf-netconf-yang-push] Clemm, A., Voit, E., Prieto, A., Tripathy, A., Nilsen-Nygaard, E., Bierman, A., and B. Lengyel, "YANG Datastore Subscription", [draft-ietf-netconf-yang-push-19](#) (work in progress), September 2018.



## [Appendix A](#). Change Log

(To be removed by RFC editor prior to publication)

v01

- o Minor revision on Subscription Decomposition
- o Revised terminologies
- o Removed most implementation related text
- o Place holder of two sections: Subscription Management, and Notifications on Subscription State Changes

v02

- o Revised [section 4](#) and 5. Moved them from appendix to the main text.

v03

- o Added a section for Terminologies.
- o Added a section for Subscription State Change Notifications.
- o Improved the Publication Composition section by adding a method to check the integrity of the data generated from different Publishers at the same time period.
- o Revised the solution overview for a more clear description.

### Authors' Addresses

Tianran Zhou  
Huawei  
156 Beiqing Rd., Haidian District  
Beijing  
China

Email: zhoutianran@huawei.com



Guangying Zheng  
Huawei  
101 Yu-Hua-Tai Software Road  
Nanjing, Jiangsu  
China

Email: zhengguangying@huawei.com

Eric Voit  
Cisco Systems  
United States of America

Email: evoit@cisco.com

Alexander Clemm  
Huawei  
2330 Central Expressway  
Santa Clara, California  
United States of America

Email: alexander.clemm@huawei.com

Andy Bierman  
YumaWorks  
United States of America

Email: andy@yumaworks.com



