Semantic Metadata Annotation for Network Anomaly Detection

draft-netana-nmop-network-anomaly-semantics-01

Experiment: Network Anomaly Lifecycle

draft-netana-nmop-network-anomaly-lifecycle-01

Helps to annotate operational data, refine outlier detection, supports supervised and semi-supervised machine learning development, enables data exchange among network operators, vendors and academia, and make anomalies for humans apprehensible

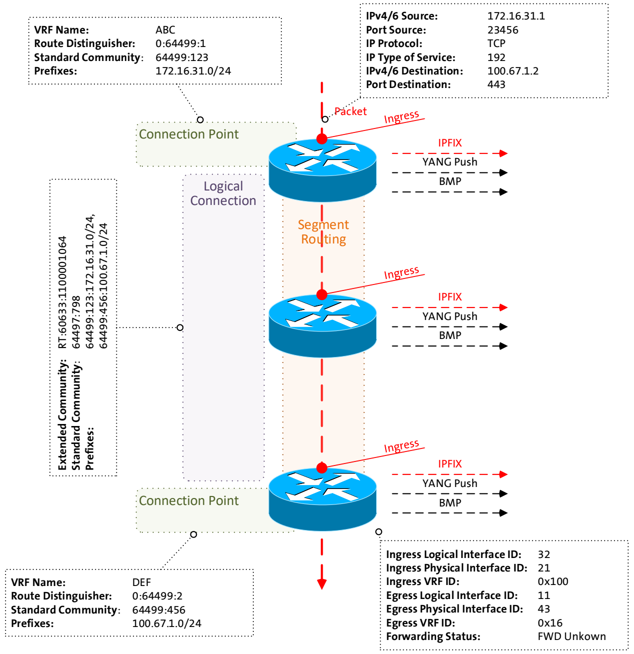
thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

12. March 2024

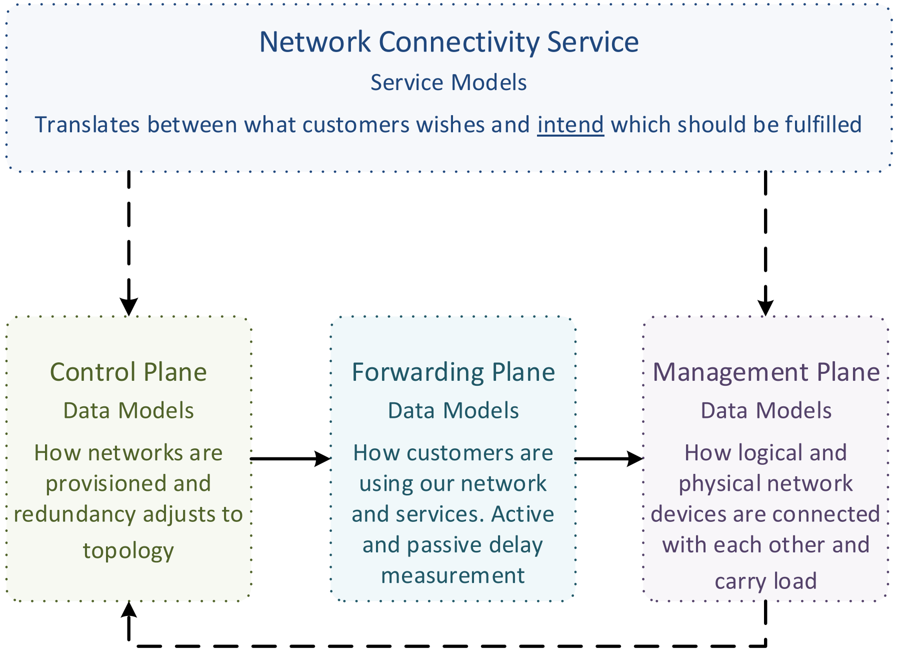
What to monitor

Which operational metrics are collected

« Network operators connect customers in routing tables called VPN's »

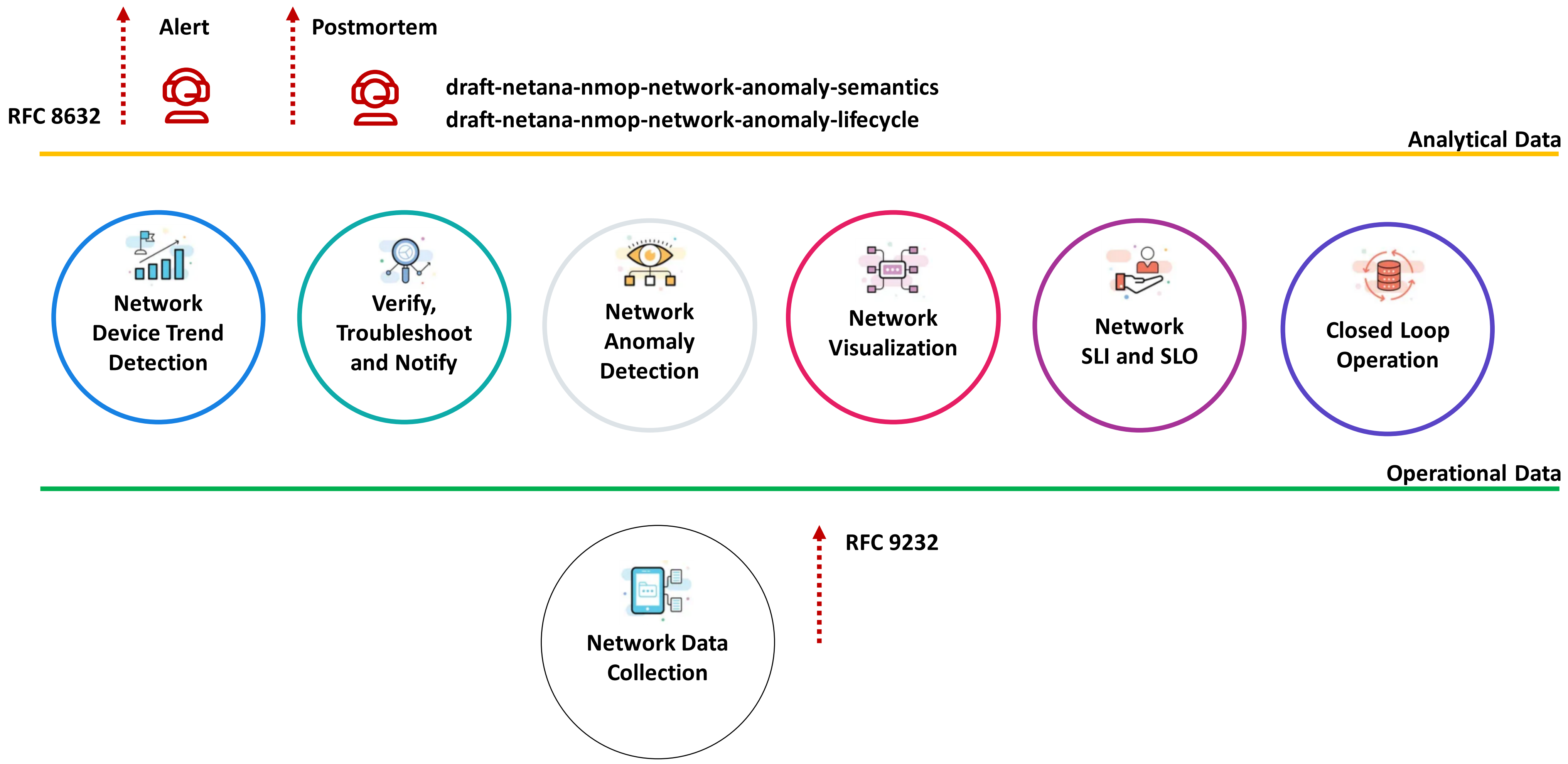


« Network Telemetry (RFC 9232) describes how to collect data from all 3 network planes efficiently »



How to organize and collaborate with data

The Data Mesh Architecture enables Network Analytics use



What our motivation is

Automate learn and improve

From network incidents postmortems we network operators **learn and improve** so does network anomaly detection and supervised and semi-supervised machine learning.

The more network incidents are observed, the more we can improve. With more incidents the **postmortem process needs be automated, let's get organized** first by defining human and machine-readable metadata semantics and annotate operational and analytical data.

Let's get further organized by exchanging standardized labeled network incident data among network operators, vendors and academia to **collaborate on academic research**.

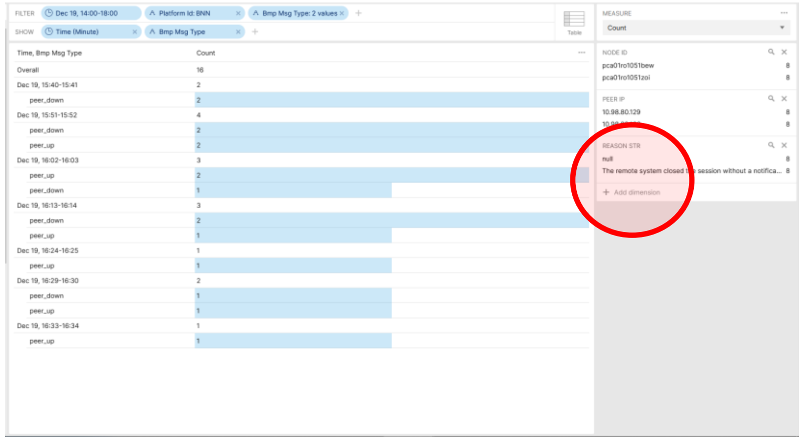
« The community working on Network Anomaly Detection is probably the only group wishing for more network incidents »

Postmortem, Maximum Prefix BGP Peer State Change

SBInfo-028166, PBI000000193943, INC000012284550

Missing Traffic 64497:6

Flow Count Drop 64497:6



BMP Peer State Change 64497:6

Traffic Drop 64497:6

IPFIX configured on PE and Inter-AS Option A ASBR nodes.

Traffic Drop with Reason Code Adjacency at TV was unrelated.

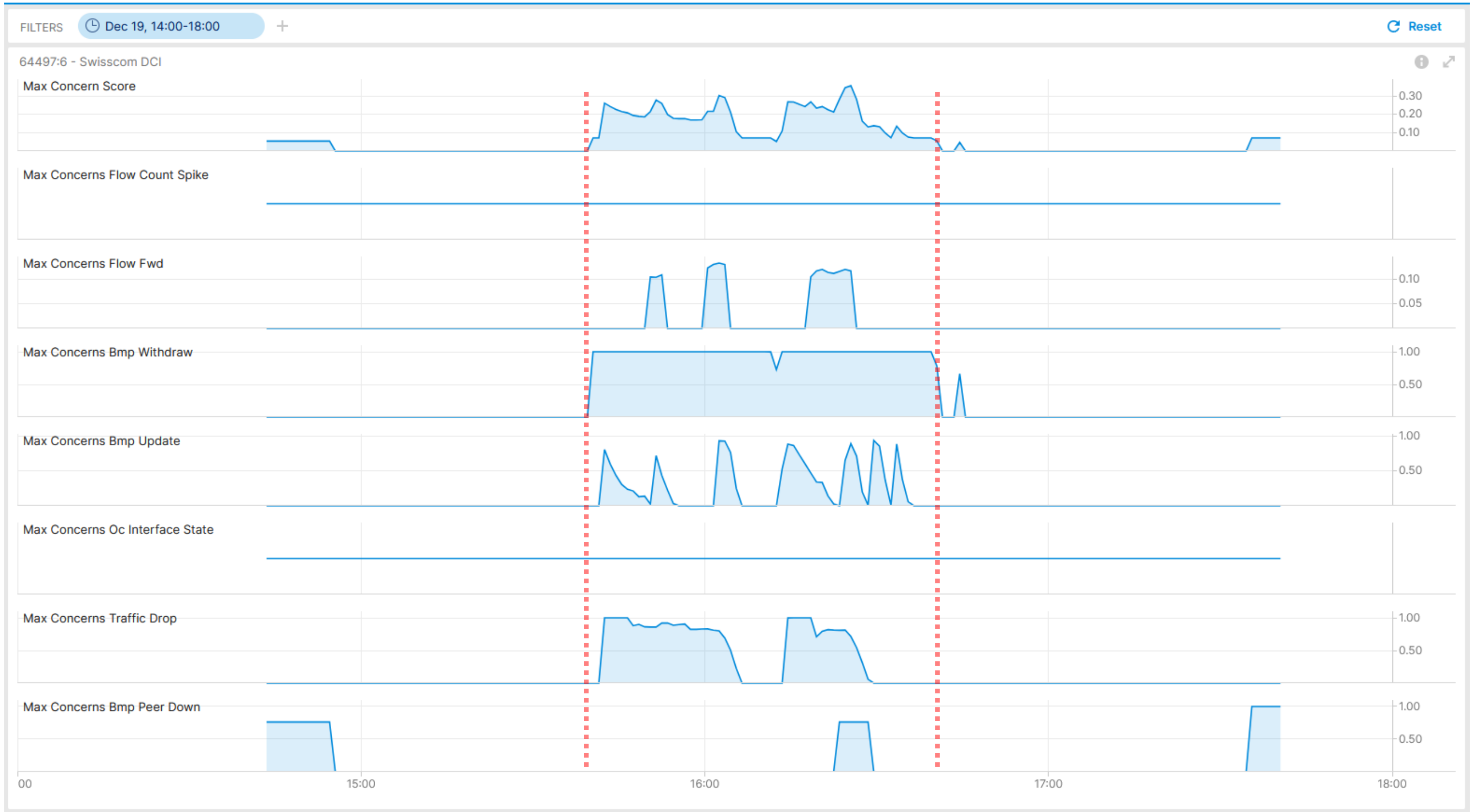
BMP ADJ-RIB In pre-policy on BGP VPNv4 /6 and IPv4/6 VRF unicast peers configured on MPLS PE's. BMP ADJ-RIB In pre-policy on BGP VPNv4 /6 on Route Reflectors.

BMP peer_down reports that it is type 4 (Remote system closed, no data) instead of type 1 (Local system closed, NOTIFICATION PDU follows) due to CSCwi61922.

Postmortem, Maximum Prefix BGP Peer State Change

SBInfo-028166, INC000012284550, Bright Lights Live

Max Concern Score: **0.36**
Traffic Drop: **1.0**
Missing Traffic: **0.13**
BMP Update/Withdraw: **1.0**
BMP Peer Down: **0.76**

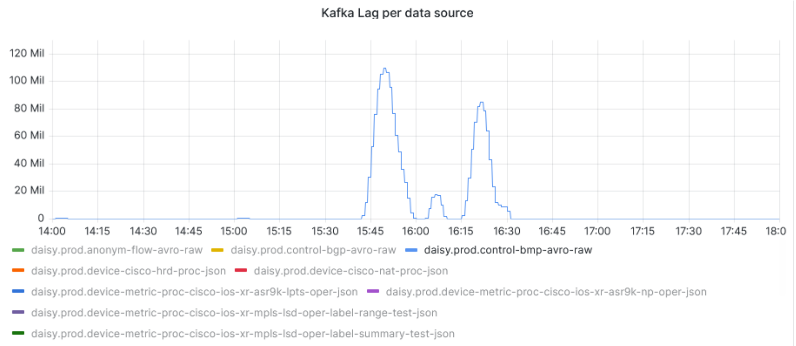


Cosmos Bright Lights Anomaly Detection – 64497:6 SC-DCI

- **BMP route-monitoring Update/Withdraw recognized topology change.**
- **BMP peer Down recognized peering state change delayed due to potential data processing lag.**
- Interface Down/Up check did not apply.
- **Traffic Drop check recognized forwarding drop.**
- **Missing Traffic recognized that connectivity is impaired.**
- Flow Count Spike did not apply.
- **Overall: 4 out of 6 checks have detected a customer impact inside of monitoring domain. Works as designed.**

Postmortem
What to do next?

- **Record incident in Cosmos Bright Lights lab. -> Done!**
- **Analyze why (TSDB ingestion delay?) not all BMP peer_down where being recognized by BMP peer_down check.**



What went well?

Anomaly Detection rules detected outage based on BMP update/withdrawal and peer_down, IPFIX flow count drop, traffic drop and missing traffic. Works as designed.

What could be improved?

Consider to implement capacity management and trend detection analytical use case for BGP max prefix configured peers, BGP Local RIB path count and BGP process memory.

draft-ietf-grow-bmp-rel authors considering to support two reason code TLV's for prefixes crossing the warning and the maximum threshold.

draft-msri-grow-bmp-bgp-rib-stats authors contacted at GROW to consider another BMP statistics definition describing how many percent of the configured maximum prefix count has been reached.

Similar as we are draft-ietf-grow-bmp-path-marking-tlv how the BGP path will be installed into the RIB, we could add as a TLV also the local allocated MPLS label from the Label FIB.

BMP peer_down reason code is 4 instead of 1 on Cisco IOS XR. Addressed and confirmed in SR 696692110. CSCwi61922 bugfix verified.

BGP notification sub-code support in NetGauze verified.

What is a symptom and how to categorize them

From action to reason to cause

Action: Which action the network node performed for a packet in the forwarding plane, a path or adjacency in the control plane or state or statistical changes in the management plane.

Reason: For each reason one or more actions describing why this action was used. From drop unreachable, administered, and corrupt in forwarding plane, to reachability withdraw and adjacency teared down in control plane, to Interface down, errors or discard in management plane.

Cause: For each reason one or more causes describes why the action was chosen. From missing next-hop and link-layer information in forwarding plane, to reachability withdrawn due to peer down or path no longer redistributed.

« Symptoms are categorized in which plane they have been observed, their action, reason and cause »

Outliers in Anomaly Detection

From global to contextual to collective

Global outliers: An outlier is considered "global" if its behavior is outside the entirety of the considered data set.

Contextual outliers: An outlier is considered "contextual" if its behavior is within a normal (expected) range, but it would not be expected based on some context. Context can be defined as a function of multiple parameters, such as time, location, etc.

Collective outliers: An outlier is considered "collective" if the behavior of each single data point that are part of the anomaly are within expected ranges (so they are not anomalous, it's either a contextual or a global sense), but the group taking all the data points together, is.

« Collective outliers are important because networks are connected. Through different planes interconnected symptoms from various angles can be observed »

Annotate Operational Data

YANG Module

```
module: ietf-symptom-semantic-metadata
  +--rw symptom
    +--rw id          yang:uuid
    +--rw event-id    yang:uuid
    +--rw description                    string
    +--rw start-timeyang:date-and-time
    +--rw end-time    yang:date-and-time
    +--rw confidence-score               float
    +--rw concern-score?                 float
    +--rw tags* [key]
    |  +--rw key        string
    |  +--rw value      string
    +--rw (pattern)?
    |  +--:(drop)
    |  |  +--rw dropempty
    |  +--:(spike)
    |  |  +--rw spike                    empty
    |  +--:(mean-shift)
    |  |  +--rw mean-shift               empty
    |  +--:(seasonality-shift)
    |  |  +--rw seasonality-shift        empty
    |  +--:(trend)
    |  |  +--rw trend                    empty
    |  +--:(other)
    |     +--rw other                    string
    +--rw source
        +--rw (source-type)
        |  +--:(human)
        |  |  +--rw human          empty
        |  +--:(algorithm)
        |     +--rw algorithm      empty
        +--rw name?                string
```

- **Symptoms** describe what changed in the network for what reason and cause with which concern score from when to when.
- **Tags** describes in which network plane, which action, reason and cause was observed.
- **Pattern** describes the measurement pattern over time of the time series data.
- **Source** describes which system **observed** the outlier. A human or a network anomaly detection system.

Experiment: Network Anomaly Post-mortem Lifecycle

What is the Motivation?

Network anomaly detection is about **identifying behaviours** that provide **evidence** of consumers experiencing a **service degradation**.

During our work to implement network anomaly detection, we realized that this requires a continuous review process, in order to **iteratively collect and incorporate more and more network and service knowledge into the methodology**.

Network operators often implement **review processes of their detection mechanisms** to improve over time (reducing both False Positives and False Negatives), **validating** the detected network anomalies and performing **post-mortem analysis**, etc.

We see the need to **provide a well-defined lifecycle for the refinement of network anomaly detection**, as this can open up to a **more structured cooperation between different actors** involved in different stages of the lifecycle, including customer service operators, network engineers, Data Scientists, AI algorithms, etc.

This proposed draft describe an **experiment**: verifying whether the approach is usable in real use case scenarios to support proper refinement and adjustments of network anomaly detection algorithms.

Network Anomaly Postmortem Lifecycle

draft-netana-nmop-network-anomaly-lifecycle

4. Lifecycle of a Network Anomaly

The lifecycle of a network anomaly can be articulated in three phases, structured as a loop: Detection, Validation, Refinement.

```
                         +-------------+
             +--------> |  Detection  | ---------+
 Adjustments |           +-------------+          | Symptoms
             |                                    |
             |                                    v
      +------------+                       +------------+
      | Refinement |<--------------------- | Validation |
      +------------+        Incident       +------------+
                          Confirmation
```

Figure 1: Anomaly Detection Refinement Lifecycle

Each of these phases can either be performed by a network expert or an algorithm or complementing each other.

Detection: The Network Anomaly Detection stage is about the continuous monitoring of the network through Network Telemetry [RFC9232] and the identification of symptoms.

Validation: Decides if the detected symptoms are signaling a real incident or if they are to be treated as false positives.

Refinement: Network operator performs detailed postmortem analysis of the network incident, collected Network Telemetry data and detected anomaly with the objective to identify useful adjustments in the Network Telemetry data collection and Anomaly Detection system.

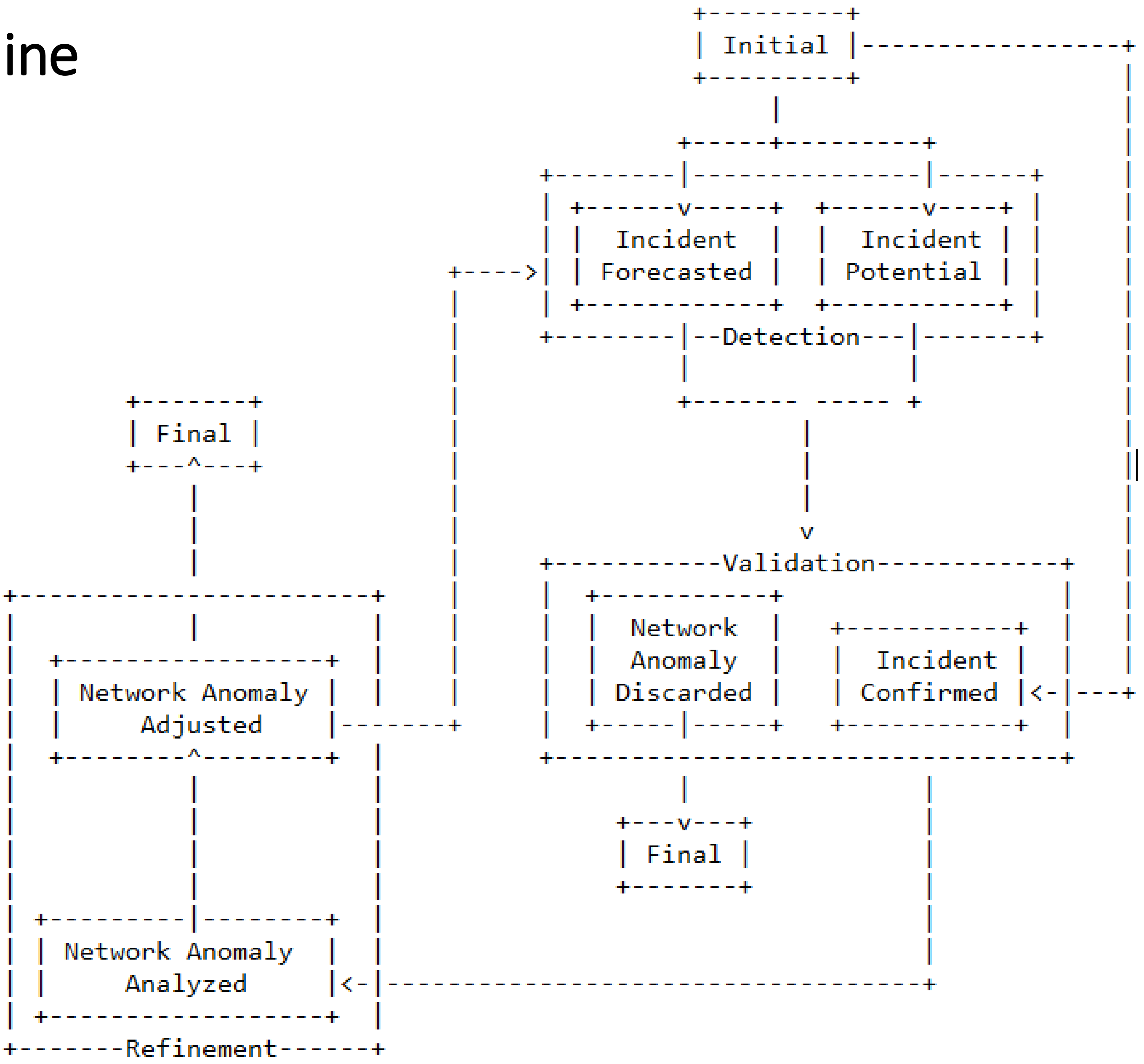
Network Anomaly State Machine

Incident Relationships

Incident Forecasted: A potential network incident is predicted in the future by the Network Anomaly Detection system.

Incident Potential: A potential network incident has been detected by the Network Anomaly Detection system.

Incident Confirmed: A potential network incident has been confirmed in the postmortem validation.



Network Anomaly Metadata

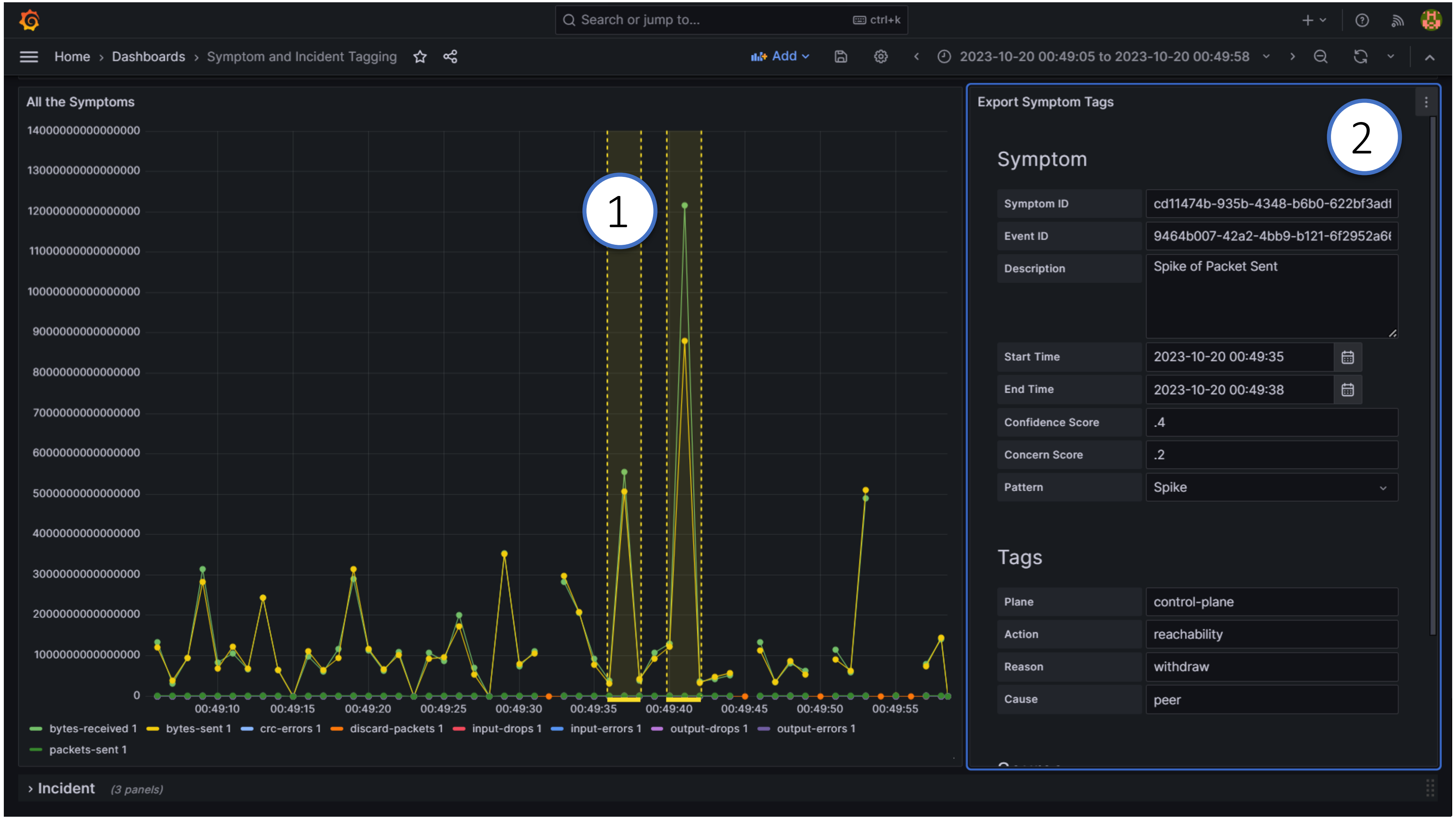
YANG Module

```
module: ietf-network-anomaly-metadata
  +--rw network-anomalies
     +--rw network-anomaly* [id author-name version state]
        +--rw id              yang:uuid
        +--rw description?    string
        +--rw author
        |  +--rw author-name      string
        |  +--rw author-type?     identityref
        |  +--rw algo-version?    uint8
        +--rw version         uint8
        +--rw state           identityref
        +--rw symptoms* [symptom_id]
           +--rw symptom_id     yang:uuid
```

- **ID and Description** uniquely identifies the detected anomaly.
- **Author Name, Type, Version and Algo-Version** describes wherever the anomaly was detected by a human or algorithm and uniquely identifies the system and version who/which detected.
- **State** describes the state of the anomaly.
- **Symptoms** describes the identified symptoms defined in ietf-symptom-semantic-metadata.

IETF 119 Hackathon – Antagonist

Labelling a Symptom in Grafana



(1) Vertical dotted lines are the tagged symptoms.

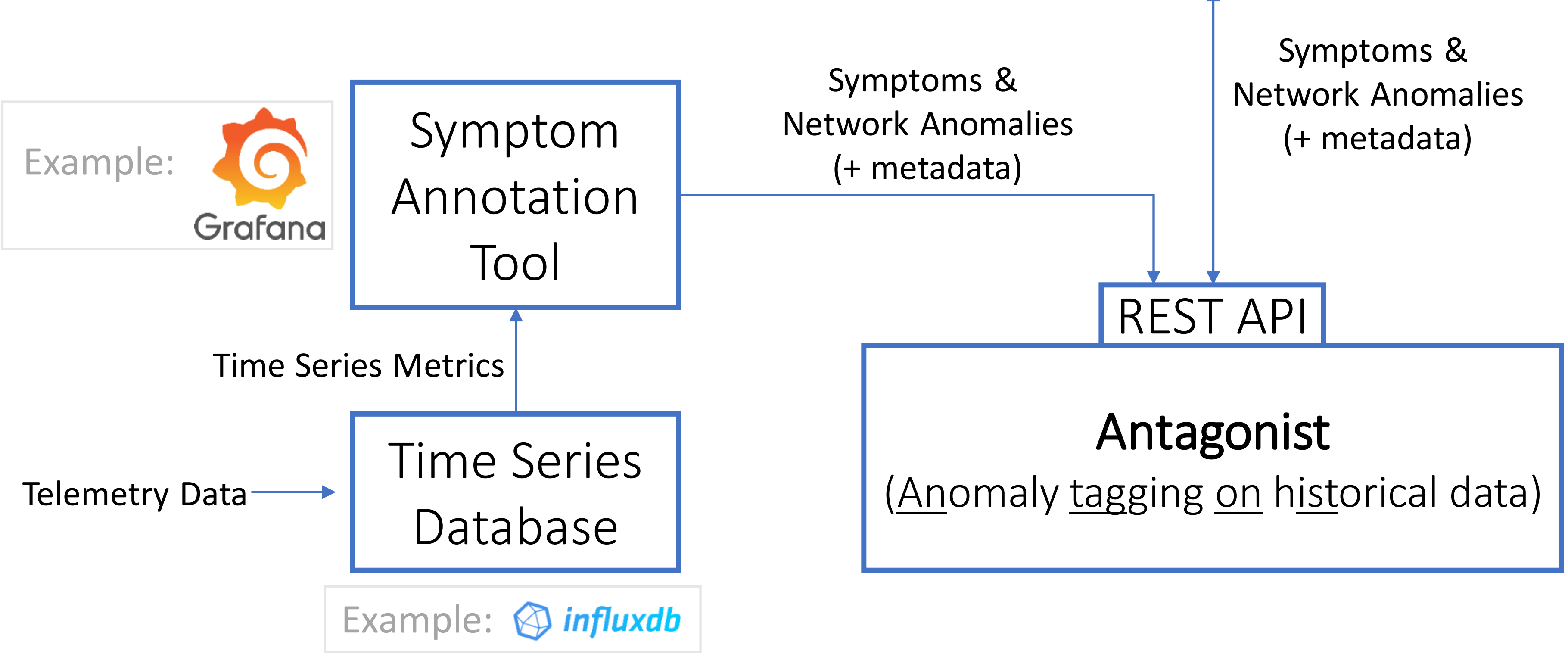
(2) Once the symptom is selected, the user can add all the details.

Once the symptom is defined it gets submitted to Antagonist.

IETF 119 Hackathon - Antagonist

Design and workflow

Antagonist has been extended to become **agnostic** of the technologies used for the persistence and annotation of telemetry data



Antagonist exposes a REST API to support **ingestion** and **exposure** of symptoms and network anomaly data and semantic metadata.

The exposed data can be used as ground-truth.

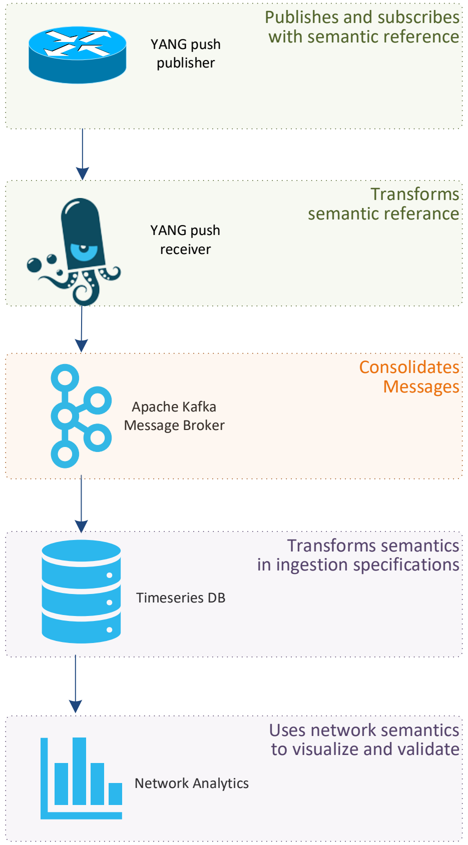
Source Code: https://github.com/vriccobene/antagonist

Add Screenshot from Antagonist Frontend

Semantic Metadata Annotation for Network Anomaly Detection

Next steps

- **This work relates to the Network Anomaly Detection topic in the NMOP charter.**
- It bridges network and data engineering, operator, vendors and academia, domains by having the **semantics and ontology of network symptoms for operational and analytical data defined**.
- This work will unveil what is missing in Network Telemetry data and provide input to other documents such as draft-davis-nmop-incident-terminology to enable a more detailed and holistic view for networks.
- **Do you realize the benefit of having standardized semantic metadata annotation for Network Anomaly Detection and how it helps network operators, vendor and academia to collaborate?**
- **-> What are your thoughts and comments?**
- **-> We request NMOP working group adoption.**



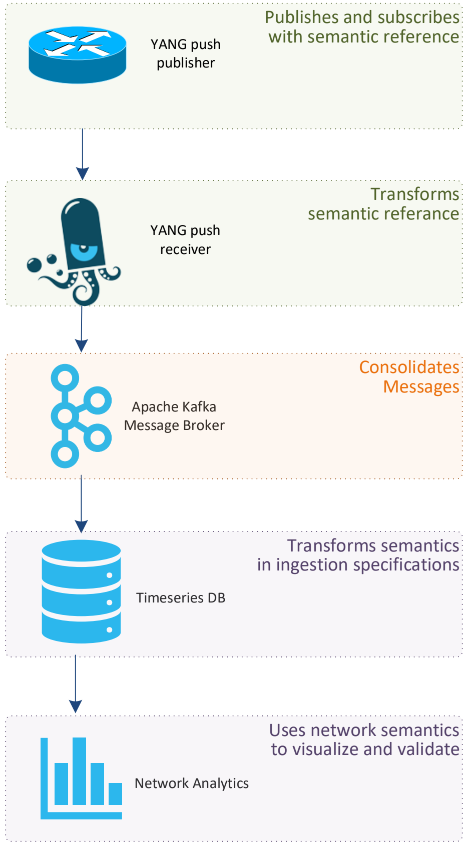
thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

12. March 2024

Experiment: Network Anomaly Lifecycle

Next steps

- **This work relates to the Network Anomaly Detection topic in the NMOP charter.**
- It defines the Network Anomaly lifecycle by providing a structured way to perform post-mortem analysis iteratively and improve the network anomaly detection methodology.
- **Our future intention: expanding and validate this approach on real use case scenarios on Swisscom network incident data.**
- This work will provide input to draft-davis-nmop-incident-terminology and complement other documents such as RFC 8632 and draft-feng-opsawg-incident-management where semantics for alerts and incidents are defined.
- **Do you realize the benefit of having a defined workflow and semantics to automate the Network Anomaly Lifecycle?**
- **-> What are your thoughts and comments?**



thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

12. March 2024