Semantic Metadata Annotation for Network Anomaly Detection

draft-netana-opsawg-nmrg-network-anomaly-semantics-01

Helps to test and validate outlier detection, supports
supervised and semi-supervised machine learning development
and make anomalies for humans apprehensible

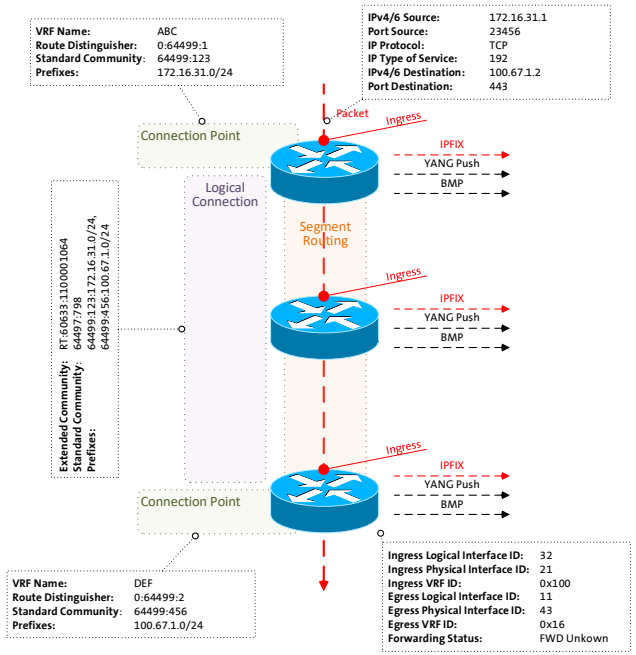
thomas.graf@swisscom.com
wanting.du@swisscom.com
alex.huang-feng@insa-lyon.fr
vincenzo.riccobene@huawei-partners.com
antonio.roberto@huawei.com

29. October 2023

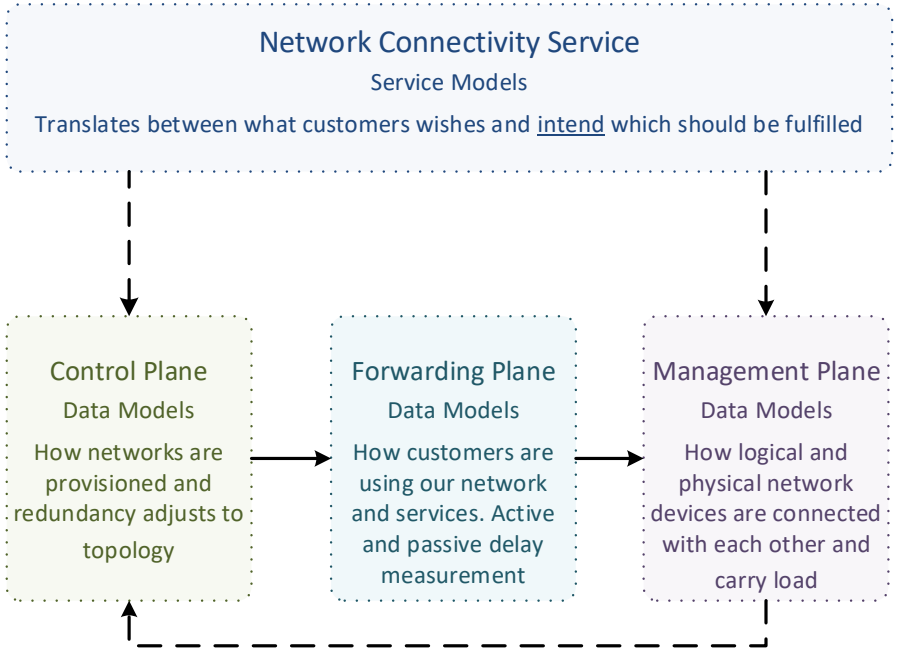
What to monitor

Which operational metrics are collected

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

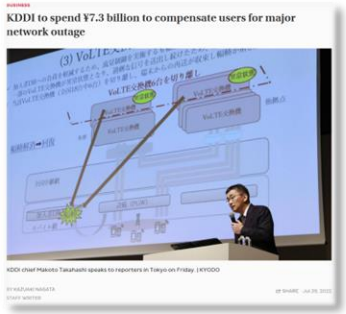


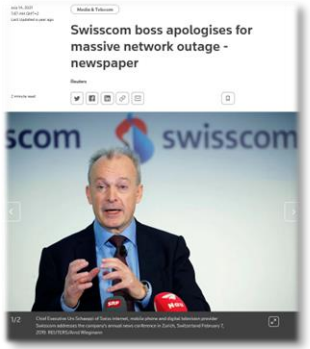
« Network Telemetry (RFC 9232) describes how to collect data from all 3 network planes efficiently »



Why to automate monitoring

Recognize network incidents faster than humans can

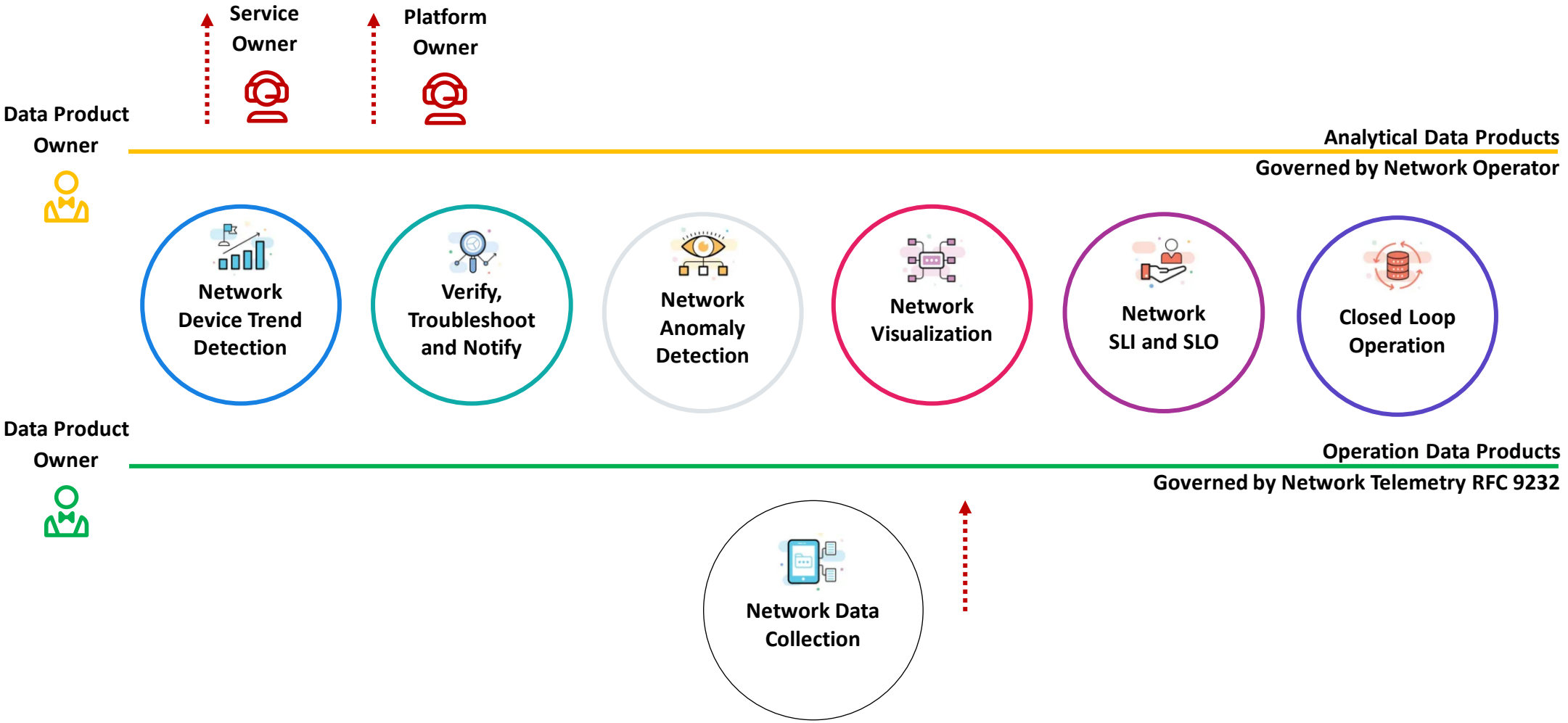




« Customers are always connected, when VPN's changing, regardless due to operational or configurational reasons, network operators are late to react due to missing visibility and automation »

How to organize and collaborate with data

The Data Mesh Architecture enables Network Analytics use



What does Network Anomaly Detection mean

Monitor changes

Network Anomaly Detection

For VPNs, Network Anomaly Detection **constantly monitors and detects any network or device topology changes**, along with their associated forwarding consequences for customers as outliers. Notifications are sent to the Network Operation Center before the customer is aware of service disruptions. **It offers operational metrics for in-depth analysis,** allowing to understand on which platform the problem originates and facilitates problem resolution.

Answers

What changed and when, on which connectivity service, and how does it impact the customers?

Focuses

Provides meaningful connectivity service impact information before customer is aware of and support in root-cause analysis.

Data Mesh

Consumes operational real-time Forwarding Plane, Control Plane and Management Plane metrics and produces analytical alerts.

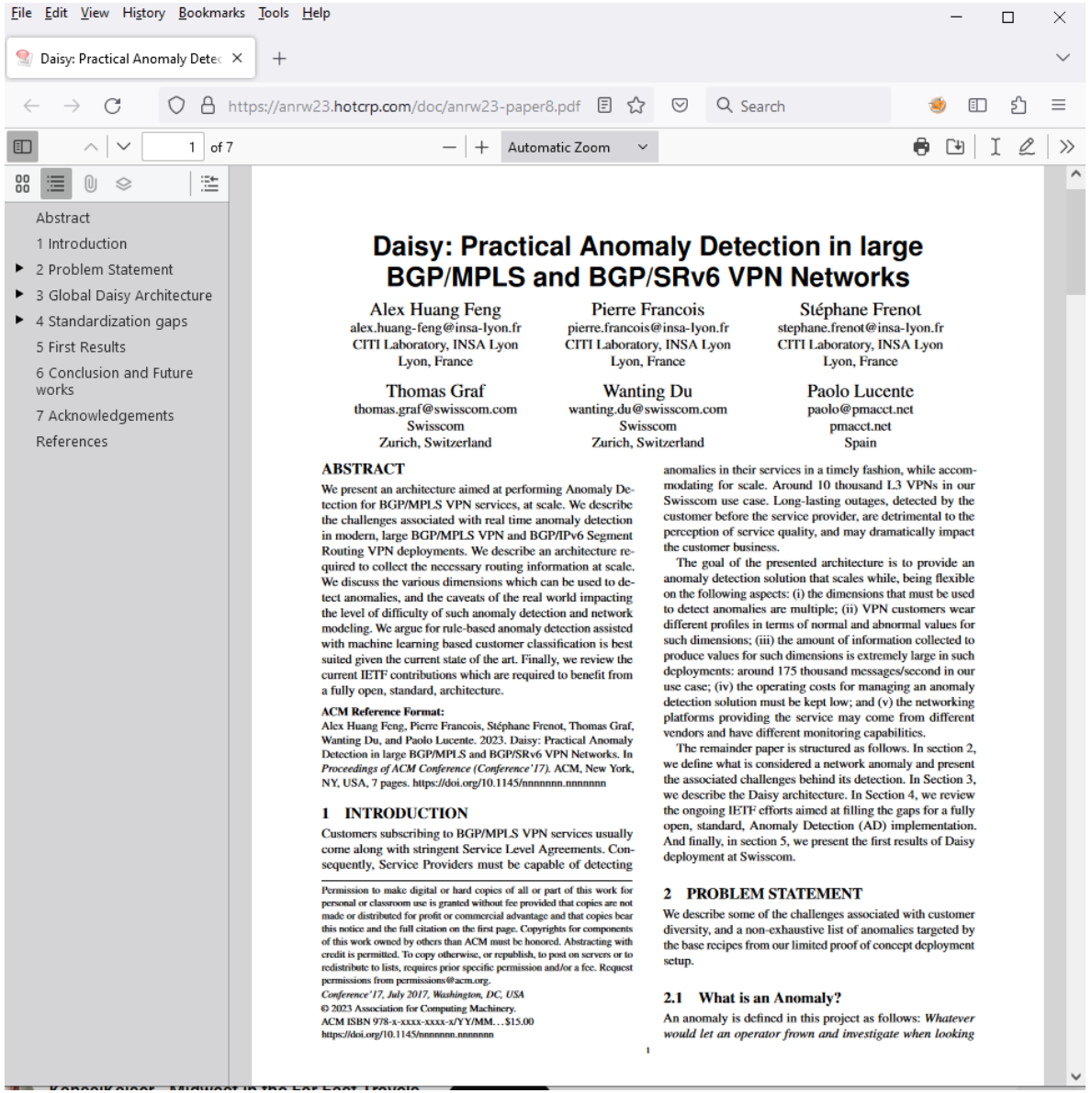
Direction

From connectivity service to network platform.

Presented in ANRW 2023

At IETF 117 San Francisco

« A more detailed paper was submitted in November 2023 to IEEE Transactions on Network and Service Management»



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 »

What is an outlier and how to categorize them

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 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 »

What is a symptom and how to categorize them

From action to reason to relation

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 action one or more reasons 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.

Relation: For each reason one or more relation describes the cause 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 »

Questions to the audience

Do you care?

Network Operators: Do you agree that today actions, traffic is dropped, path is withdrawn, interface is down are always exposed through Network Telemetry, but reason and cause, dropped due to unreachable next-hop, withdrawn due to peer down, interface down due to missing signal, are rarely and would be most interesting?

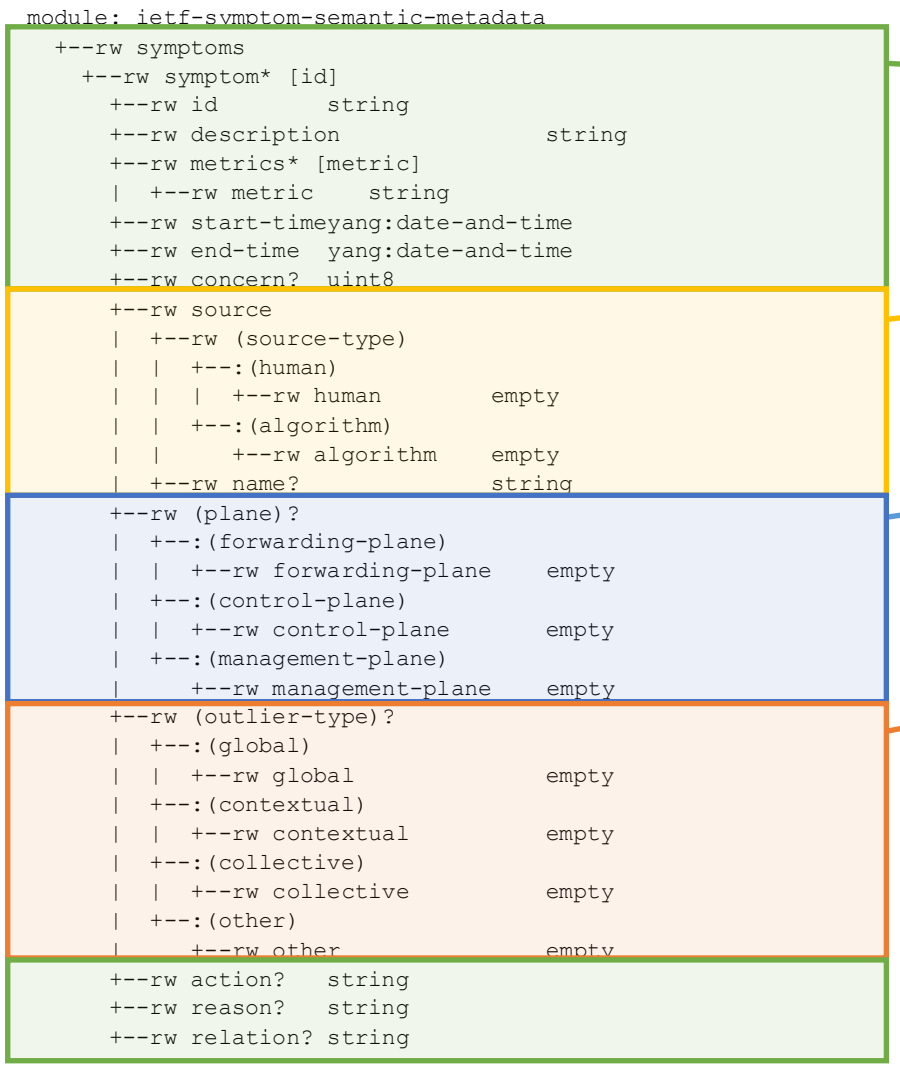
Network Vendors: Is the assumption correct that a network process, routing process withdrawing a path, most of the time knows why it acts that way, and could potential make this reason and cause information available?

Academia: Would it help if network operators would provide well defined labeled operational and analytical data to enable and validate their research?

Everybody: Should these symptoms be clearly described and standardized for a common terminology so that operators, researchers and anomaly detection systems alike understand their meaning and learn and act accordingly?

Annotate Operation Data

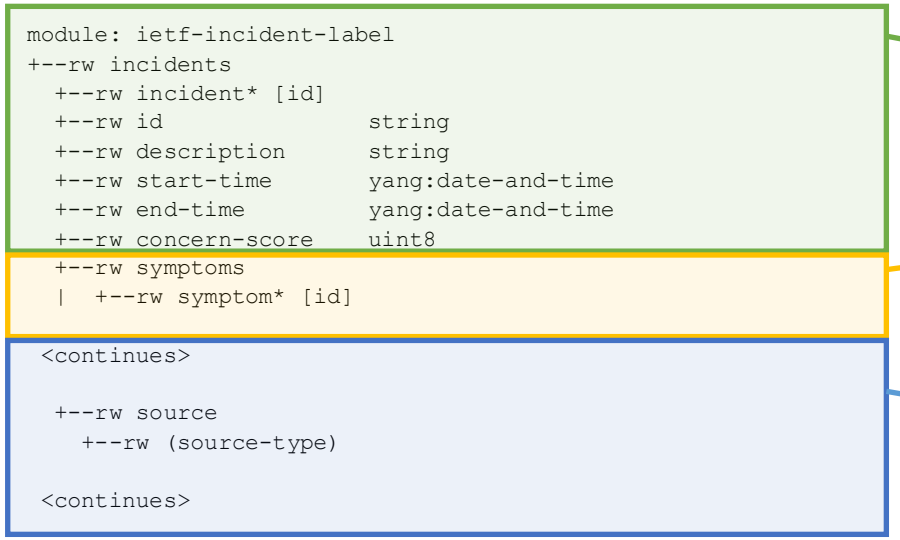
YANG Module



- **Symptoms** describe what changed in the network for what reason and cause with which concern score from when to when.
- **Source** describes who detected the outlier. A human or a network anomaly detection system.
- **Plane** describes in which network plane it was observed. Forwarding, Control or Management Plane.
- **Outlier-Type** describes which type of outlier it is. Global, Contextual or Collective.

Annotate Analytical Data

YANG Module

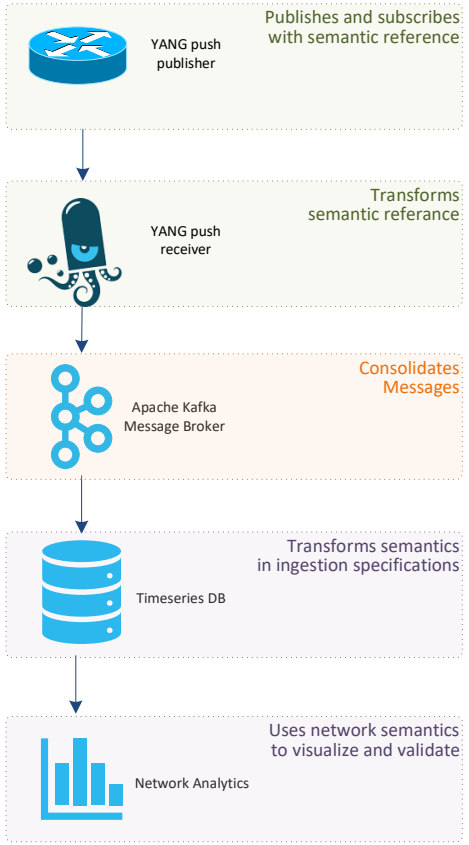


- **Incidents** has a unique ID and description with a start and end time and a concern score.
- **Symptoms** describe what changed in the network for what reason and cause with which concern score from when to when.
- **Source** describes who detected the outlier. A human or a network anomaly detection system.

Semantic Metadata Annotation for Network Anomaly Detection

Next steps

- **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?**
- This document looks for a community and working group who have interest in Network Anomaly Detection, bridging network and data engineering, operator, vendors and academia, by writing the **semantics and ontology of network symptoms for operational and analytical data**.
- This work will unveil what is missing in Network Telemetry data and provide input for other documents to enable a more detailed and holistic view from networks.



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

29. October 2023