An Architecture for a Network Anomaly Detection Framework

draft-netana-nmop-network-anomaly-architecture-00

Motivation and architecture of a Network Anomaly Detection Framework
and the relationships to other documents describing
network symptom semantics and network incident lifecycle

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

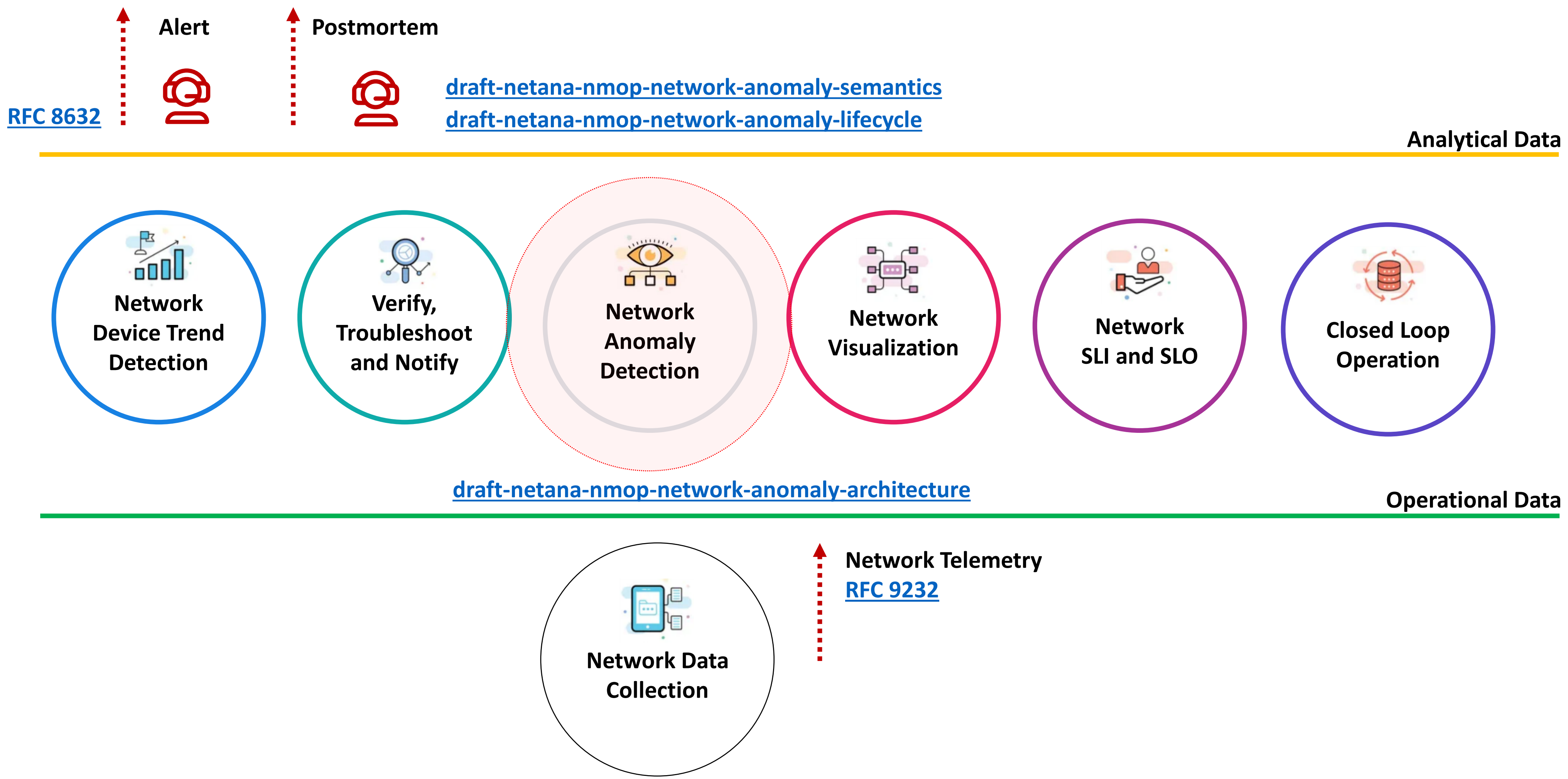
25. July 2024

Why This I-D?
A Reminder

- This document describes motivation and a generic and extensible architecture of a Network Anomaly Detection Framework.
- Anchors draft-netana-nmop-network-anomaly-semantics and draft-netana-nmop-network-anomaly-lifecycle documents.
- Different applications will be described and exampled with open-source running code.

Structuring Anomaly Detection NMOP Effort

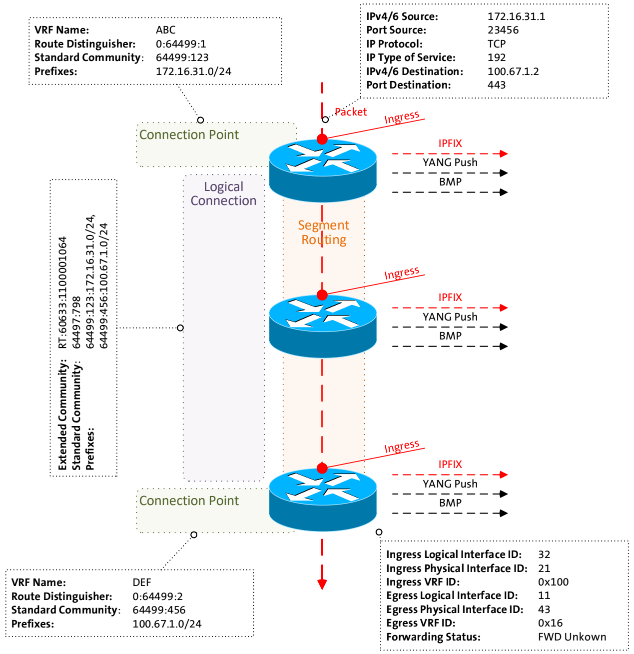
Integrates into Data Mesh



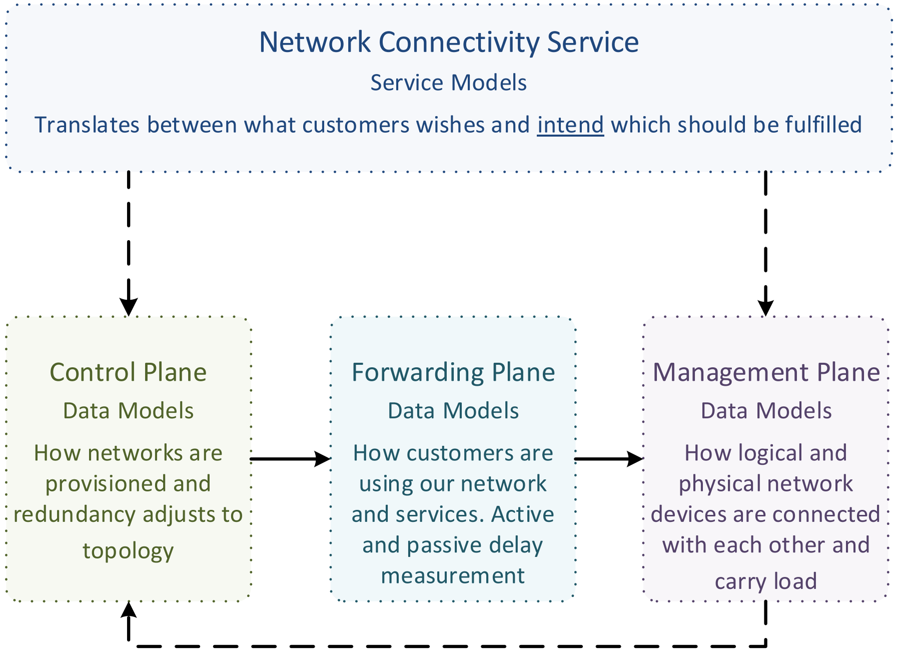
What to monitor

Which metrics are collected

« Network operators connect customers in routing tables called Connectivity Services »



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



What does Network Anomaly Detection mean

Monitor changes, called outliers, in networks

Network Anomaly Detection

For Connectivity Services, Network Anomaly Detection **constantly monitors and detects any network or device topology change**, 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 in 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.

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 »

Elements of the Architecture

```
        +----------+                    +--------------------+
        |Service   |                    |     Alert and      |
   |--- |Inventory |                    | Problem Management |
   |    |          |                    |       System       |
   |    +----------+                    +--------------------+
   |      |                                   ^     Stream
   |      |                                   |
   |      |         +---------+         +--------------------+
   |      |         | Post-   | Stream  |   Message Broker   |
   |      |         | mortem  | <-------|  with Analytical   |
   |      |         | System  |         |    Network Data    |
   |      |         +---------+         +--------------------+
   |      |              |                    ^     Stream
   |      |              |                    |
   |      |              |              +--------------------+
   |      | Profile      | Fine         | Alert Aggregation  | Store Label
   |      | and          | Tune         | for Anomaly        | ------------|
   |      | Generate     | SDD          | Detection          |             |
   |      | SDD Config   | Config       +--------------------+             |
   |      |              |                  ^  ^  ^ Stream                 |
   |      v              v                  |  |  |                        ?
   |   +--------------------+           +--------------------+      +----------+
   |   | Service Disruption| Schedule   | Service Disruption | Replay | Data     |
   |   |     Detection     | ---------->|     Detection      |<------ | Storage  |
   |   |   Configuration   | Detection  |                    |        |          |
   |   +--------------------+           +--------------------+        +----------+
   |                                      ^ ^ Stream ^ ^ ^                 ^
   |                                      | |        | | |                 |
   |                                    +---------+----------+             |
   |                                    | Network | Data     | Store       |
   |----------------------------------->| Model   | Aggr.    | ------------|
   |                                    |         | Process  | Operational Data
   |                                    +---------+----------+
                                            ^  ^  ^ Stream
                                            |  |  |
                                        +--------------------+
                                        |   Message Broker   |
                                        |  with Operational  |
                                        |    Network Data    |
                                        +--------------------+
                                            ^  ^  ^ Stream
Subscribe                       Publish     |  |  |
          +--------------------+        +--------------------+
          | Network Node with  | ------>| Network Telemetry  |
--------->| Network Telemetry  | ------>|  Data Collection   |
          |    Subscription    | ------>|                    |
          +--------------------+        +--------------------+
```

- **Service Inventory** contains list of the connectivity services.
- **Service Disruption Detection** processes aggregated network data to decide whether a service is degraded or not.
- **Service Disruption Detection Configuration** defines the set of approaches that need to be applied to perform SDD.
- **Operational Data Collection** manages network telemetry subscriptions and transforms data into message broker.
- **Operational Data Aggregation** produces data upon which detection of a service disruption can be performed.
- **Network Modeling** establishes knowledge of network relationships.
- **Data Profiling** categorizes nondeterministic customer related data.
- **Detection Strategies** for a profile a detection strategy is defined.
- **Machine Learning** is commonly used to detect outliers or anomalies.
- **Storage** some algorithms may relay on historical (aggregated) operational data to detect anomalies.
- **Alerting** consolidates analytical insights and notifies.
- **Postmortem** refines and stores the network anomaly and symptom labels into the Label Store.
- **Replaying** to validate refined anomaly and symptom labels, historical operational data is replayed.

Semantic Metadata Annotation for Network Anomaly Detection

draft-netana-nmop-network-anomaly-semantics

```
module: ietf-symptom-semantic-metadata
  +--rw symptom
     +--rw id?                              yang:uuid
     +--rw event-id?                        yang:uuid
     +--rw description?                     string
     +--rw start-time?                      yang:date-and-time
     +--rw end-time?                        yang:date-and-time
     +--rw confidence-score?                score
     +--rw concern-score?                   score
     +--rw tags* [key]
     |  +--rw key       string
     |  +--rw value     string
     +--rw (pattern)?
     |  +--:(drop)
     |  |  +--rw drop                       empty
     |  +--:(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 annotator
        +--rw (annotator-type)
        |  +--:(human)
        |  |  +--rw human          empty
        |  +--:(algorithm)
        |     +--rw algorithm      empty
        +--rw name?                string
```

- **Symptom ID and description** uniquely identifies the detected anomaly. **Event ID, start/end-time and confidence/concern-score** uniquely identifies the network event with its start and end time, how confident the system identified the anomaly and how concerned an operator should be.
- **Tags** allows to add customer information.
- **Pattern** describes the identified pattern of the anomaly.
- **Annotator Name, Type,** describes wherever the anomaly was detected by a human or algorithm and uniquely identifies the system who/which detected.

Experiment: Network Anomaly Lifecycle

draft-netana-nmop-network-anomaly-lifecycle

« Network Anomaly Detection is an iterative process that requires continuous improvement »

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.

Experiment: Network Anomaly Lifecycle

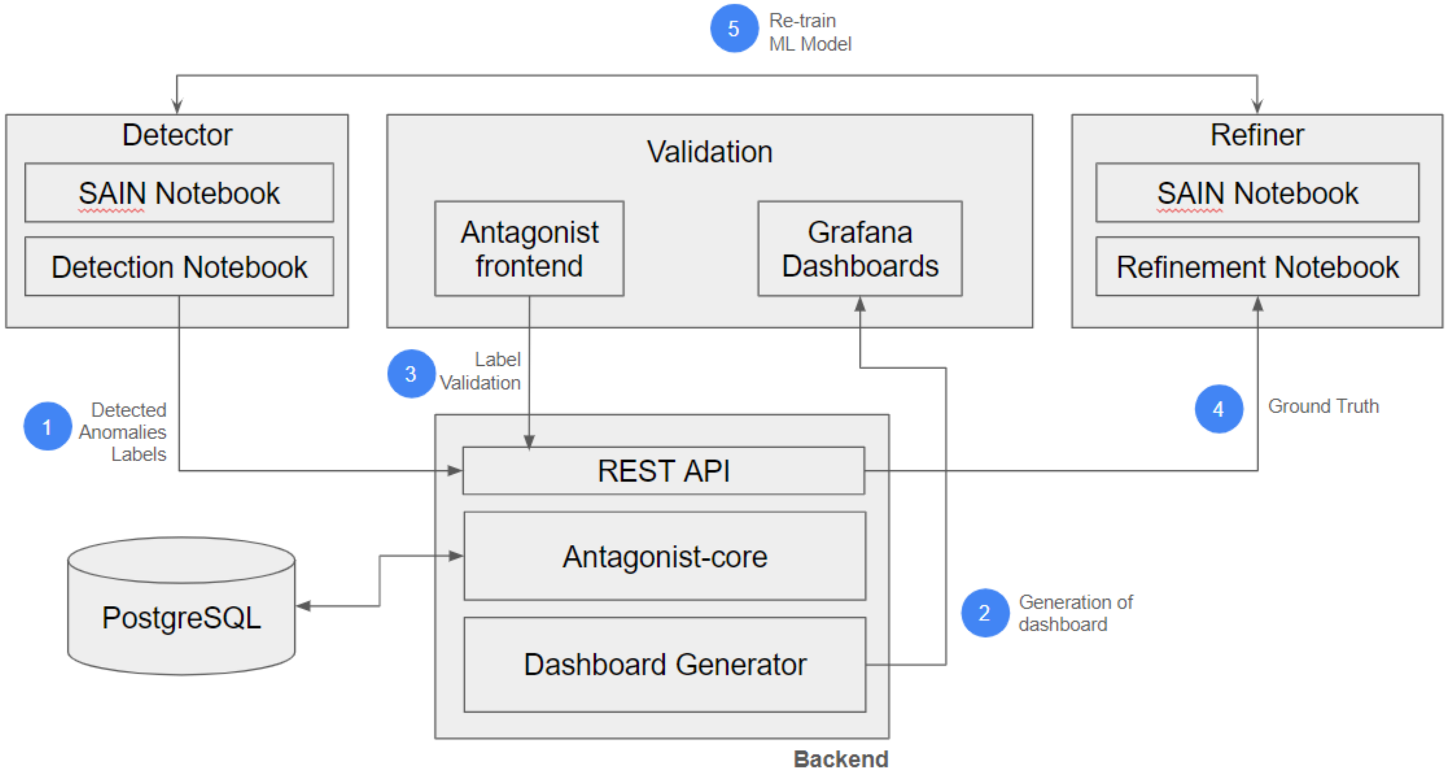
draft-netana-nmop-network-anomaly-lifecycle

```
module: ietf-network-anomaly-metadata
  +--rw network-anomalies
    +--rw network-anomaly* [id version]
      +--rw id                yang:uuid
      +--rw version           uint32
      +--rw description?      string
      +--rw state             identityref
      +--rw annotator
      |  +--rw (annotator-type)
      |  |  +--:(human)
      |  |  |  +--rw human          empty
      |  |  +--:(algorithm)
      |  |     +--rw algorithm      empty
      |  +--rw name?                empty
      +--rw symptoms* [symptom_id]
        +--rw symptom_id     yang:uuid
```

- **ID and Description** uniquely identifies the detected network anomaly (as a container of symptoms).
- **Description and State** provide general information regarding the anomaly and .
- **Annotator** describes the entity that observed the network anomaly: this can be a human or an algorithm (anomaly detection system).
- **Symptoms** provides a list of symptoms (based on ietf-symptom-metadata) that are part of this network anomaly.

Experiment: Antagonist

anomaly tagging on historical data



Next Steps:

- Improve scalability
- Validate with Swisscom Data

Goals:

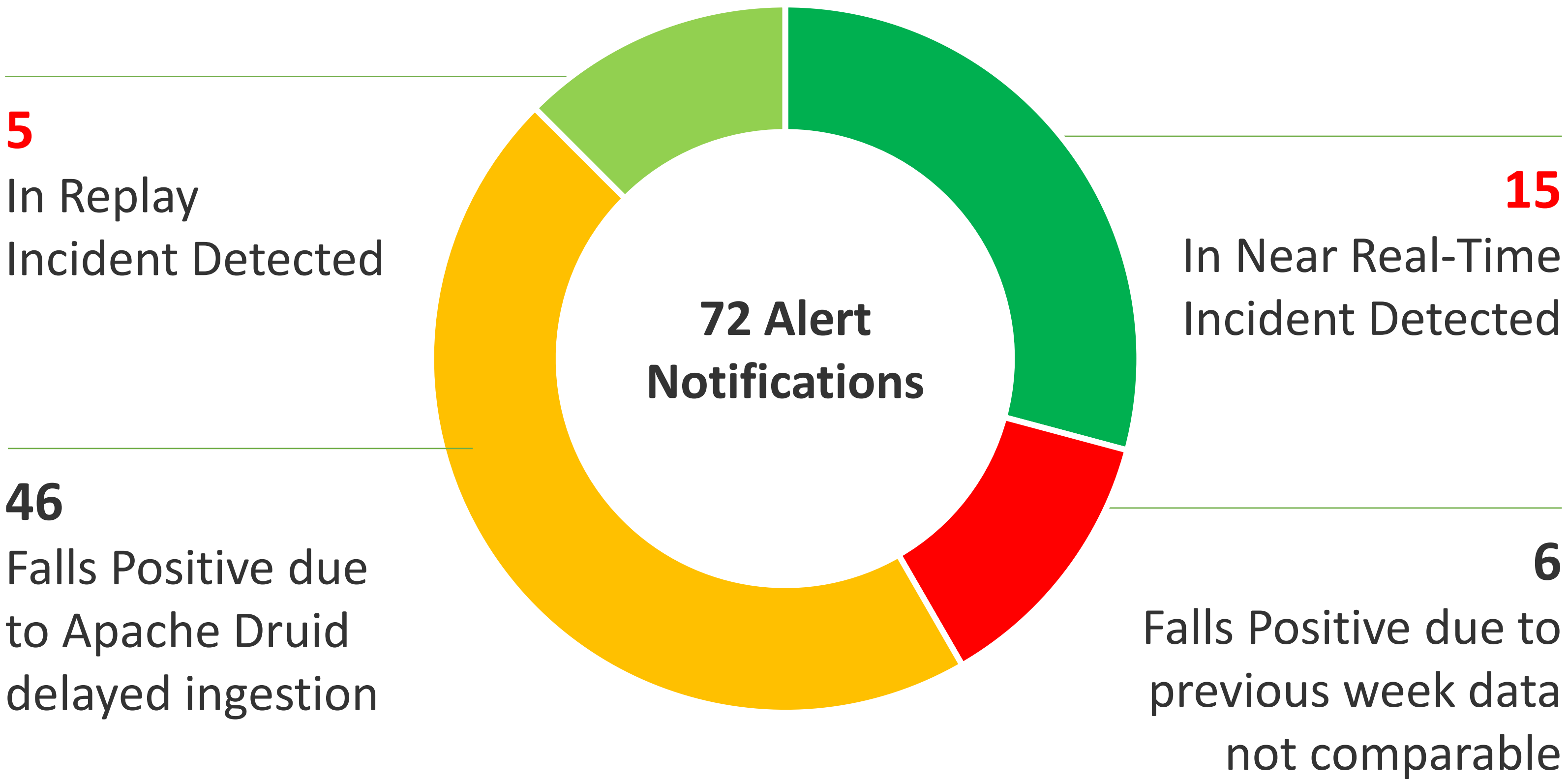
- Prove that YANG models contain all the necessary information
- Validate models across a wide range of use-cases
- Show interoperability between

Done so far:

- ✓ Validation with real operational data (Cloud monitoring)
- ✓ Validation with rule-based Network Anomaly Detector (SAIN RFC9417/RFC9418)
- ✓ Validation with a ML-based Network Anomaly Detector (Autoencoder)
- ✓ Add support for Re-training of ML-based models
- ✓ Add partial support for Metadata Filtering and search
- ✓ YANG model refinements to reflect the results of the coding
- ✓ Automatic dashboard generation

Swisscom - Cosmos Bright Lights PoC Summary

After 20 Incidents and 18 Months Time



Key Facts in V0 (2023-2024)

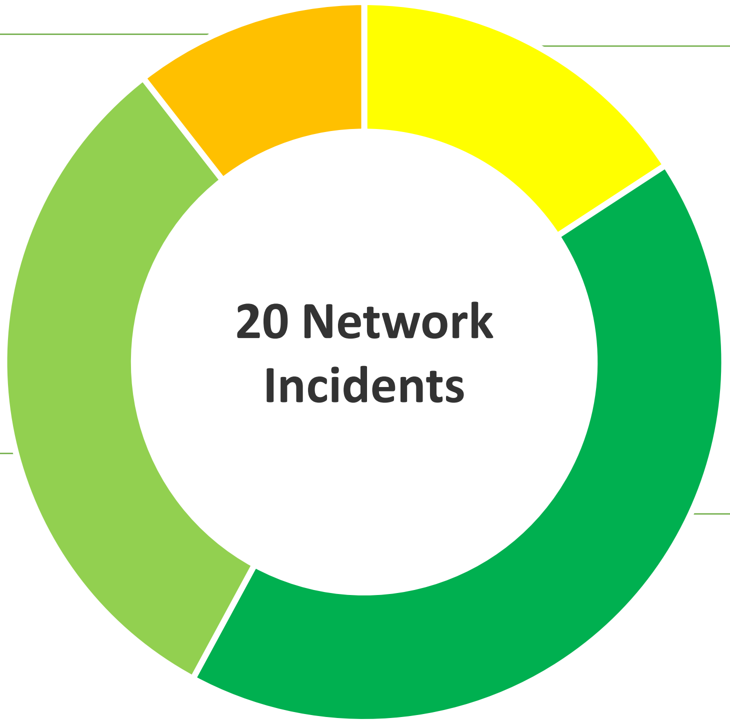
- 16 L3 VPNs proactively monitored.
- Individual Service Disruption Detection rule accuracy is beyond 90%. Summed accuracy is beyond 95%.
- Max Concern score ranged between 0.06 and 0.85. In average 0.46.
- In 4 cases additional YANG, in 13 cases additional BMP, in 2 cases Netconf Transaction-ID and 1 case additional L2 IPFIX metrics would have helped to gain more visibility.
- Key observability feature missing: BMP Local RIB with Path Marking.

Swisscom – PoC Detail and Outlook

Multiple Perspectives increases Accuracy

3
One out of Six checks

6
Three out of Six checks



2
Four out of Six checks

8
Two out of Six checks

Key Improvements in V1 (2024)

- >12000 L3 VPNs proactively monitored since June 2024.
- Realtime Streaming eliminates delayed ingestion falls positives and scaling.
- Improved profiling. Compares to multiple previous weeks and discard largest deviation eliminates falls positives.
  -> Work In progress

Key Improvements in V2 (2025)

- Annotate operational and analytical Network Incident data for reproduction.
- Enabling automated workflow. From PowerPoint slide decks to data driven actionable insights.

An Architecture for a Network Anomaly Detection Framework

Status, Summary and Next steps

Status of draft-netana-nmop-network-anomaly-architecture

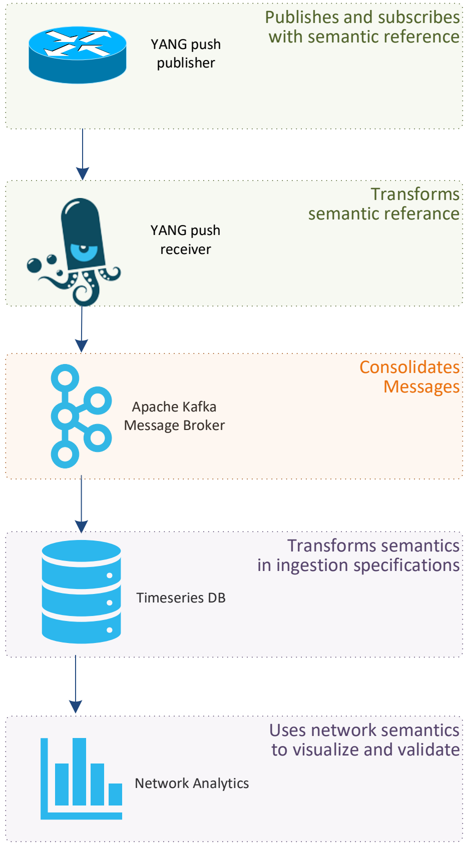
- Reference document to anchor anomaly detection work items.

Status of draft-netana-nmop-network-anomaly-semantics and draft-netana-nmop-network-anomaly-lifecycle

- Referenced draft-netana-nmop-network-anomaly-architecture as the architecture document.
- Change the term source to annotator and updated the YANG modules accordingly.
- Added/updated terminology section with references to draft-ietf-nmop-terminology and draft-netana-nmop-network-anomaly-architecture.
- Moved data mesh and outlier detection section to draft-netana-nmop-network-anomaly-architecture.

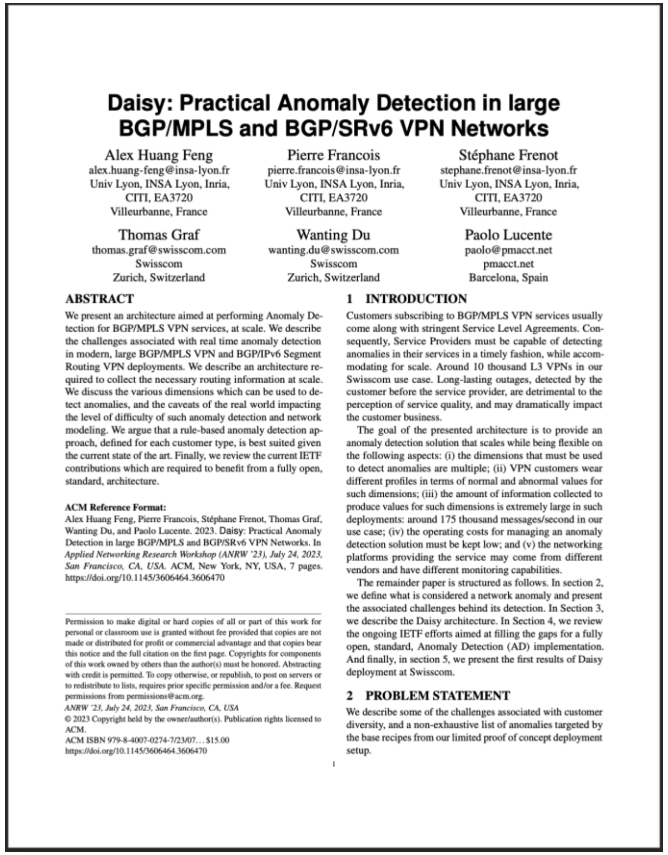
Next Steps

- ➢ **Request adoption for all 3 anomaly detection documents starting with draft-netana-nmop-network-anomaly-architecture.**
- ➢ **NMOP interim meeting on September 11th proposal**
  - ➢ **Network incident postmortem examples from Swisscom and Bell Canada**
  - ➢ **Detailing documents, updates and hackathon experiment results**
  - ➢ **Invite other operators to contribute on experiments**



Relevant Papers for more Details

Paper "**Practical Anomaly Detection in Internet Services: An ISP centric approach**"
Published at AnNet Workshop (In conjunction with IEEE NOMS)
Seoul, South Korea (6–10 May 2024)
Open access: https://hal.science/hal-04655324



Paper "**Daisy: Practical Anomaly Detection in large BGP/MPLS and BGP/SRv6 VPN Networks**" published at ACM/IRTF ANRW'23
San Francisco, USA (24 July 2023)
Open access: http://hal.science/hal-04307611