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

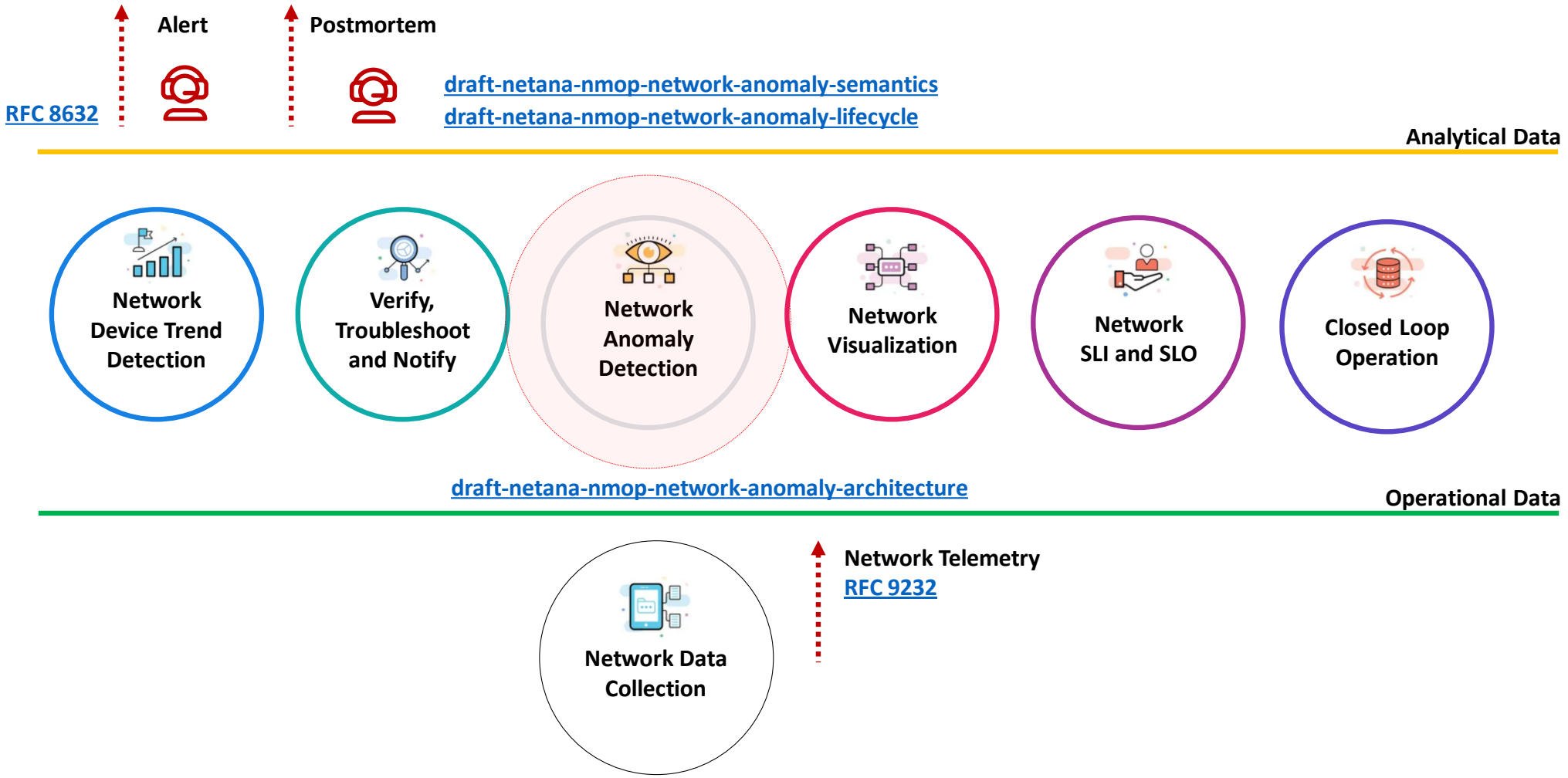
Why This I-D?

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

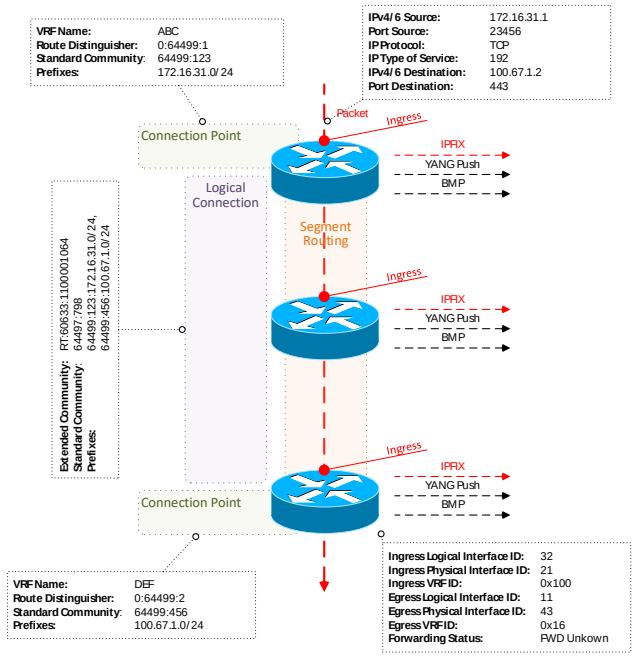
Integrates into Data Mesh



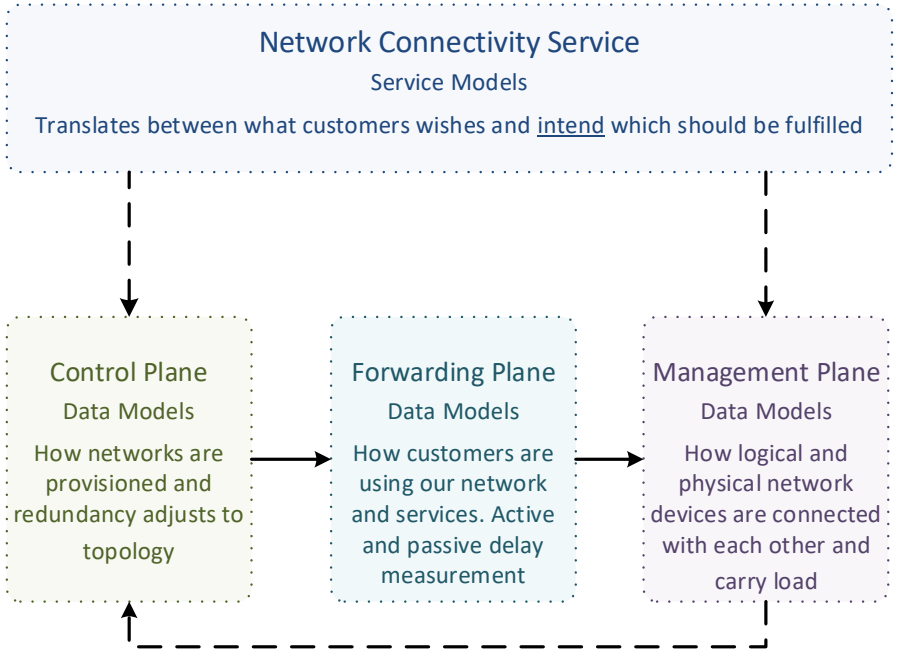
What to monitor

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

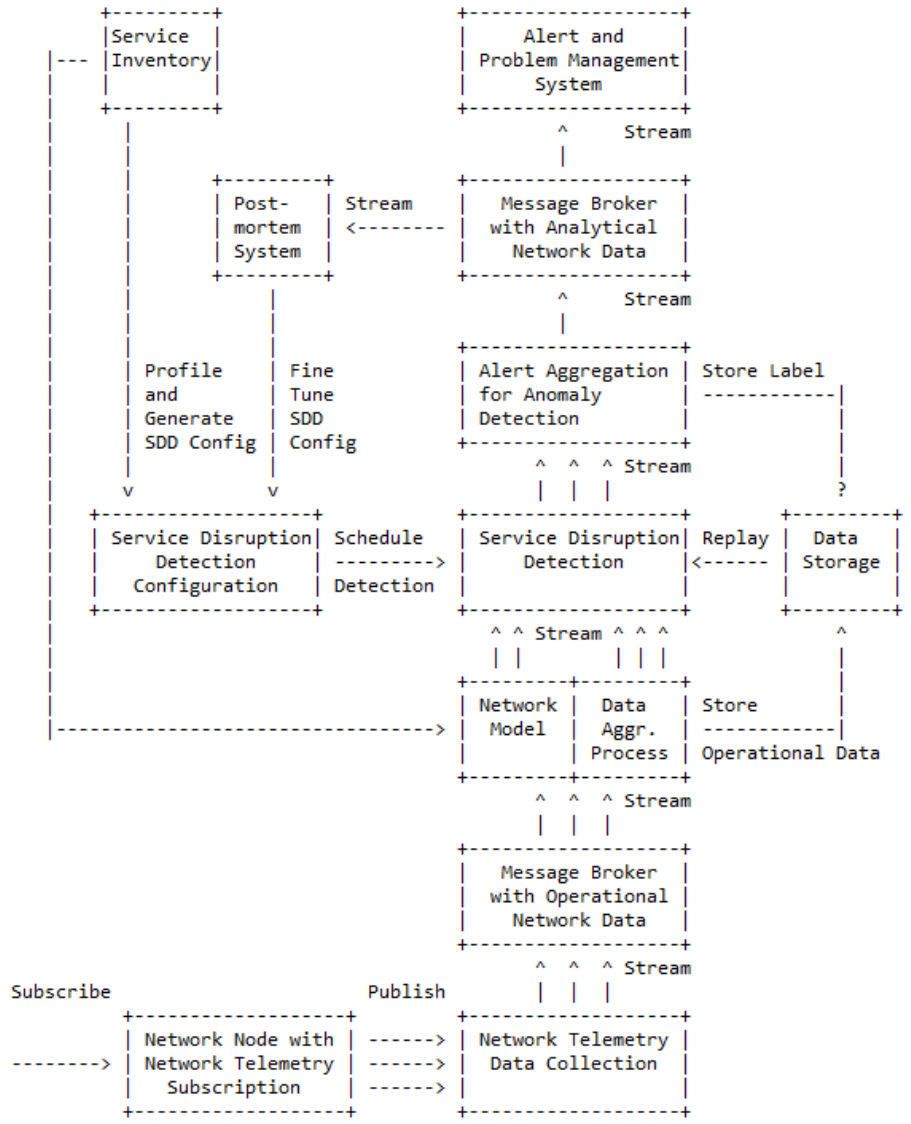
From network incidents postmortems we network operators **learn and improve** so does network anomaly detection and supervised and semi-supervised machine learning.

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 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

Goal: Enable the exchange of labelled dataset for network anomaly detection between operators, vendors and academia

```
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 symptom with its start and end time, how confident the system identified the anomaly and how concerned an operator should be.
- **Tags** describe the semantic metadata of the symptom).
- **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 entity 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.

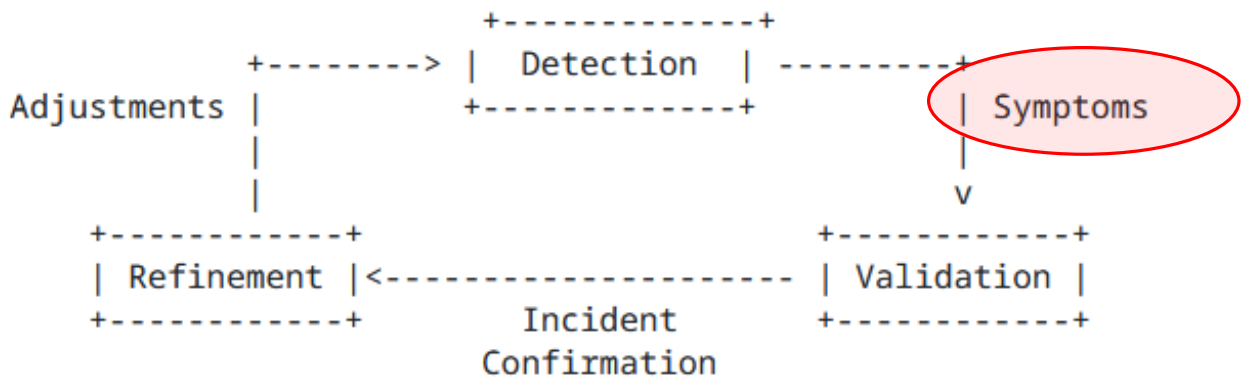


Figure 1: Anomaly Detection Refinement Lifecycle

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

This draft defines:

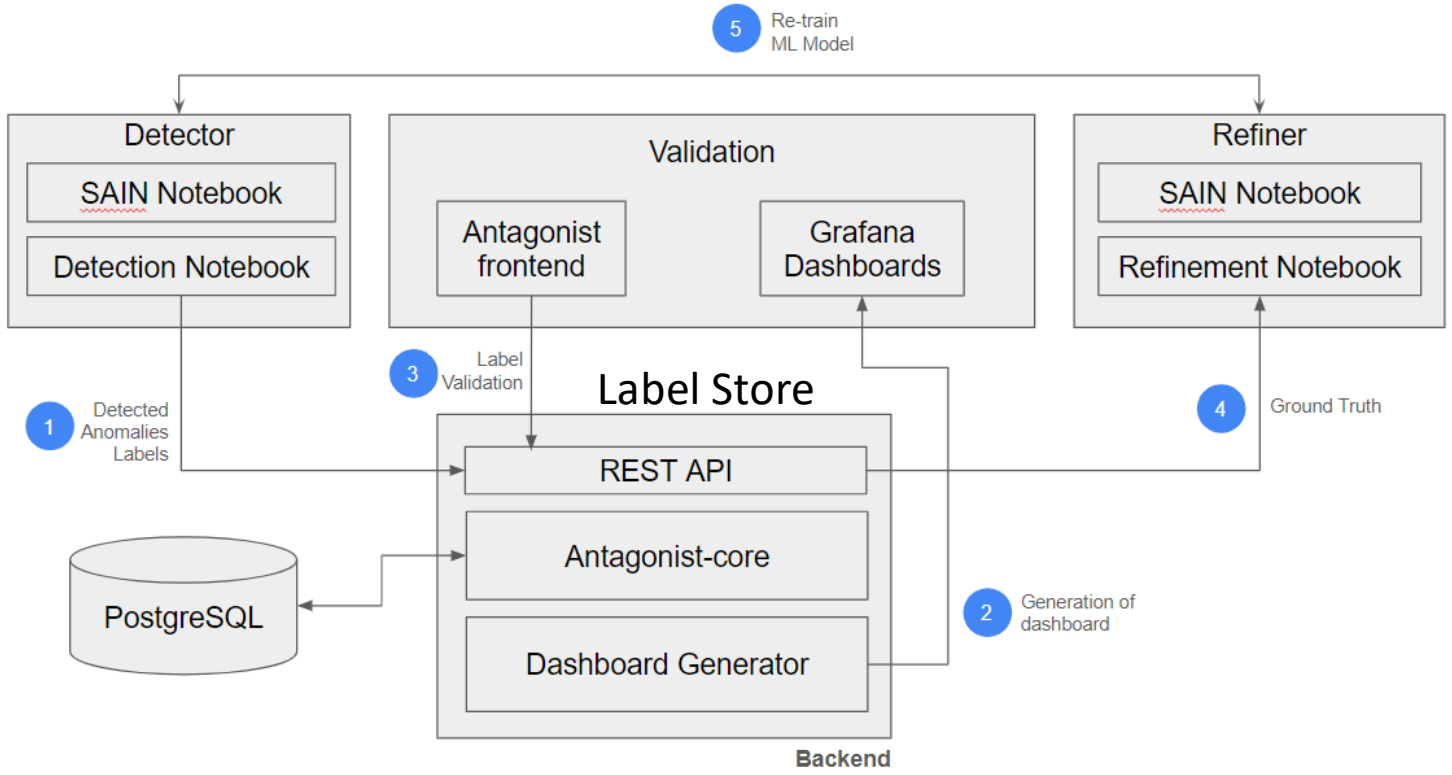
- A **State machine for network anomalies** spanning across the whole lifecycle
- A **YANG data model** describing the network anomaly as a collection of symptoms

```
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 its current state in the lifecycle.
- **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 that are part of this network anomaly.

Experiment: Antagonist

anomaly tagging on historical data



Next Steps:

- Improve scalability
- Refine YANG Models
- Integrate and 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 stages

Done so far:

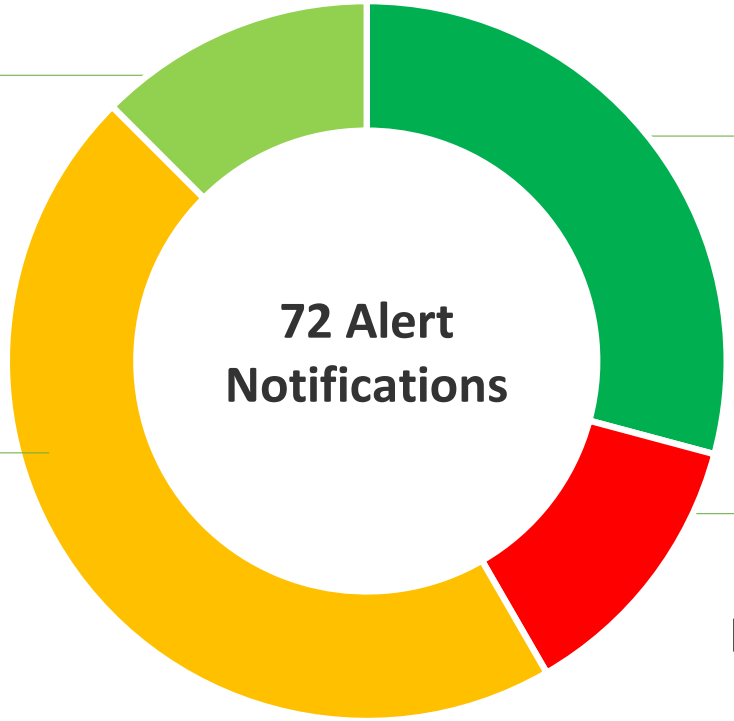
- ✓ Validation with **real operational data** (Cloud monitoring)
- ✓ Validating with rule-based Network Anomaly Detector (**SAIN RFC9417/RFC9418**)
- ✓ Validation with a **ML-based Network Anomaly Detector** (Autoencoder)
- ✓ Add support for **Re-training** of ML-based models
- ✓ Add partial support 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

5
In Replay
Incident Detected

46
Falls Positive due to Apache Druid delayed ingestion



15
In Near Real-Time
Incident Detected

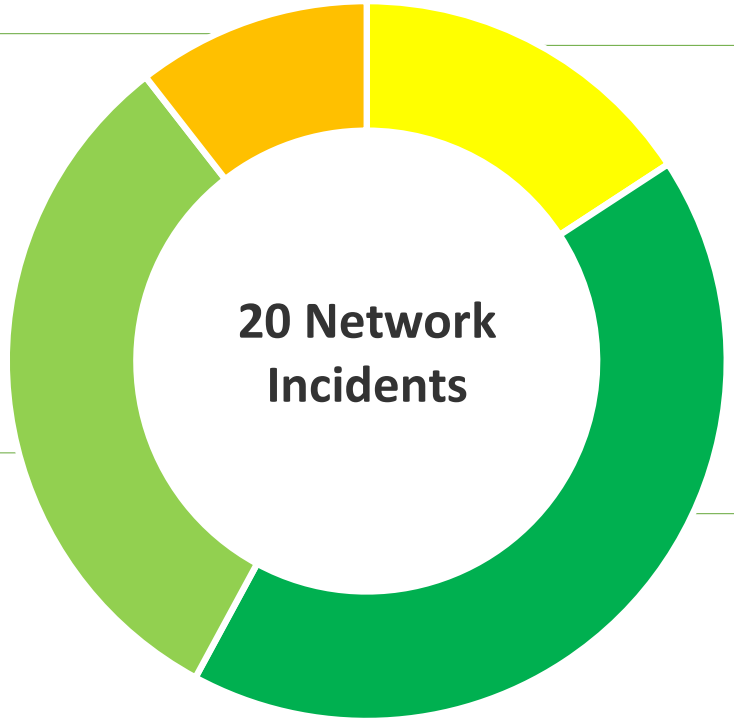
6
Falls Positive due to previous week data not comparable

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

2
Four out of Six checks

6
Three 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

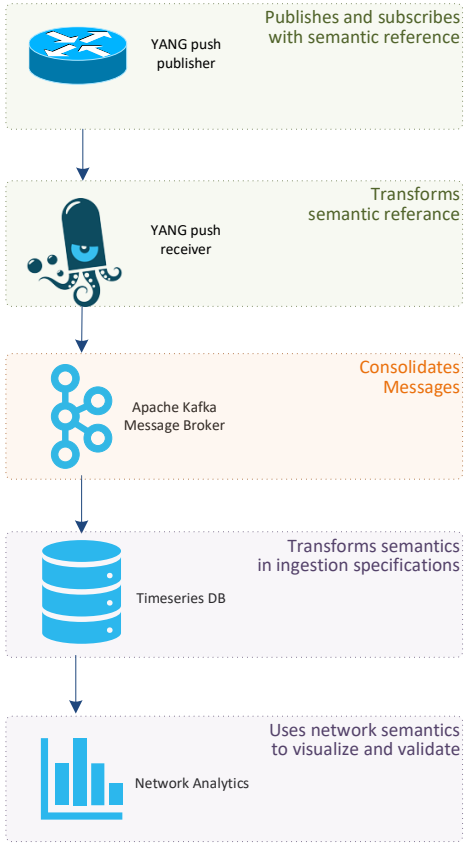
- Reference document to anchor anomaly detection work items.

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

Practical Anomaly Detection in Internet Services: An ISP centric approach

Alex Huang Feng*, Pierre Francois*, Kensuke Fukuda‡, Wanting Du†, Thomas Graf†, Paolo Lucente§, Stéphane Frénot*

*INSA Lyon, Inria, CITI, UR3720, Villeurbanne, France
alex.huang-feng@insa-lyon.fr, pierre.francois@insa-lyon.fr, stephane.frenot@insa-lyon.fr
‡National Institute of Informatics, Tokyo, Japan
kensuke@nii.ac.jp
†Swisscom, Zurich, Switzerland
wanting.du@swisscom.com, thomas.graf@swisscom.com
§pmacct.net, Barcelona, Spain
paolo@pmacct.net

***Abstract*—Identifying anomalies in a network is a crutial endeavor for Internet Service Providers (ISPs). Anomalies that impact the traffic of the ISP customers can lead to a degradation in the reputation of the company. Moreover, silent anomalies that do not break connectivity can impact the revenue and business of ISPs. Therefore, monitoring and anomaly detection has become essential for ISPs. In this paper, we present an ongoing research project aimed at identifying anomalies in Internet services provided by an ISP. We aim at detecting anomalies within the domain managed by the ISP that impact the customers and the business of the ISP. We propose a rule-based approach designed to promptly detect and provide reporting for such anomalies in near real time, giving information that allows the operator to identify whether a solution can be brought. In this paper, we describe the collected network telemetry metrics and illustrate how they are processed using open-source solutions. We introduce a set of use cases showing that an ISP can monitor Internet services using IETF standard metrics.**

I. Introduction

Internet services include providing global Internet reachability for customer Autonomous Systems (ASes) connected to an Internet Service Provider (ISP) and serving private customers within the ISP (e.g. FTTH). Disruptions in the network that affect the connectivity of an ISP not only significantly degrade the organization's reputation but also have implications on the company's revenue. Customers subscribed to Internet services depend on the ISP peerings to reach the Internet and an incident between them and the Internet can have detrimental implications for their business.

Today, routing between different ASes is established using BGP [1]. ISPs managing an AS configure policies in their routers based on the business relationship they have with their neighboring ASes. Generally, ISPs classify their BGP neighbors into Customers, Settlement-free Peers and Transit Providers. Customer ASes compensate the ISP to reach the Internet, Settlement-free peers are mutual arrangements between two ISPs to exchange Internet traffic without any financial compensation and Transit Providers provide access to the global Internet.

ISPs rely on collected BGP messages and traffic counters to monitor peerings and detect anomalies that could impact their customers. They closely supervise network traffic to identify unexpected patterns or potential abuses by peers. This underscores the importance for ISPs to receive prompt alerts when anomalous or unwanted traffic behaviors occur, enabling network operators to rapidly implement solutions and address the detected issues.

Anomaly detection (AD) has been a hot topic in the last decade where researchers have proposed new ways to detect irregularities in the data. Most research projects aiming at detecting anomalies in BGP networks use public repositories such as Routeviews and RIPE NCC archives [2, 3], allowing researchers to identify problems in Internet from a global point of view. In conjunction with publicly known incidents, researchers have been able to develop methods to detect anomalies in data from the public domain, with a focus on detecting anomalies in the global Internet topology [4, 5].

Simulated environments mimicking the deployed network and manually generated anomalies have also been used to test anomaly detection [6]. Very few projects use production data coming from an ISP to detect anomalies and root cause analysis within a single domain. AD within an AS have only been investigated by very few researchers having access to production data [7]–[9].

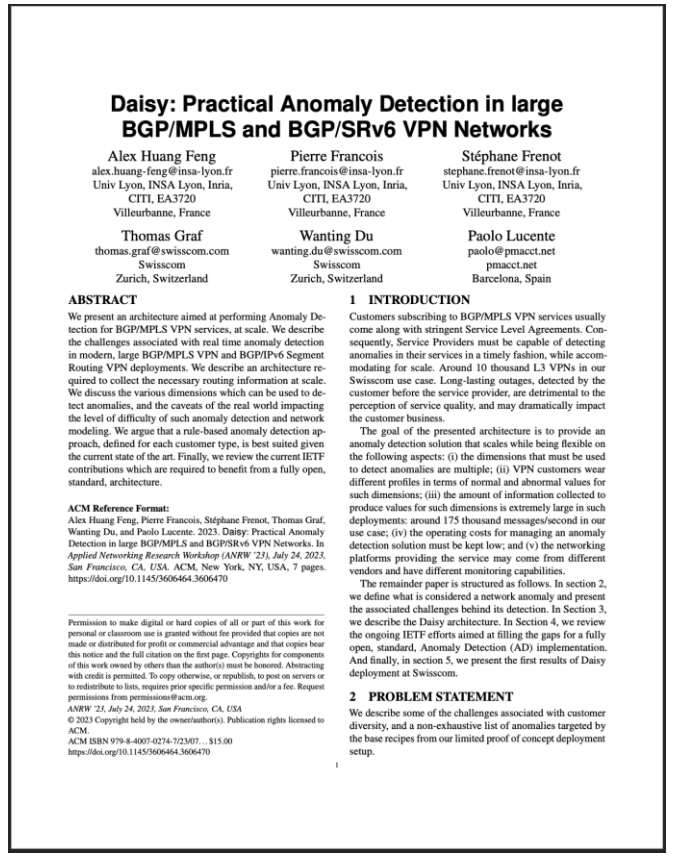
In this paper, we focus on detecting anomalies within a single AS to potentially help them fixing their configuration and find unwanted traffic flows impacting their business. We describe the target use cases in Section II. Instead of solely using BGP activity as a source of data, as done in [7], we use a larger set of monitoring information, allowing us to cover a broader set of service anomalies (Sec. III-A). The authors in [8] focus on detecting performance issues from end-to-end users, while the work presented in this paper also covers anomalies impacting the traffic from peerings. In [9], anomaly detection is based on traffic information with a focus on network intrusion detection, while the project presented

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



Daisy: Practical Anomaly Detection in large BGP/MPLS and BGP/SRv6 VPN Networks

Alex Huang Feng
alex.huang-feng@insa-lyon.fr
Univ Lyon, INSA Lyon, Inria, CITI, EA3720
Villeurbanne, France

Pierre Francois
pierre.francois@insa-lyon.fr
Univ Lyon, INSA Lyon, Inria, CITI, EA3720
Villeurbanne, France

Stéphane Frenot
stephane.frenot@insa-lyon.fr
Univ Lyon, INSA Lyon, Inria, CITI, EA3720
Villeurbanne, France

Thomas Graf
thomas.graf@swisscom.com
Swisscom
Zurich, Switzerland

Wanting Du
wanting.du@swisscom.com
Swisscom
Zurich, Switzerland

Paolo Lucente
paolo@pmacct.net
pmacct.net
Barcelona, Spain

ABSTRACT

We present an architecture aimed at performing Anomaly Detection for BGP/MPLS VPN services, at scale. We describe the challenges associated with real time anomaly detection in modern, large BGP/MPLS VPN and BGP/IPv6 Segment Routing VPN deployments. We describe an architecture required to collect the necessary routing information at scale. We discuss the various dimensions which can be used to detect anomalies, and the caveats of the real world impacting the level of difficulty of such anomaly detection and network modeling. We argue that a rule-based anomaly detection approach, defined for each customer type, is best suited given the current state of the art. Finally, we review the current IETF contributions which are required to benefit from a fully open, standard, architecture.

ACM Reference Format:
Alex Huang Feng, Pierre Francois, Stéphane Frenot, Thomas Graf, Wanting Du, and Paolo Lucente. 2023. Daisy: Practical Anomaly Detection in large BGP/MPLS and BGP/SRv6 VPN Networks. In *Applied Networking Research Workshop (ANRW '23), July 24, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3606464.3606470

ANRW '23, July 24, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0274-7/23/07…$15.00
https://doi.org/10.1145/3606464.3606470

1 INTRODUCTION

Customers subscribing to BGP/MPLS VPN services usually come along with stringent Service Level Agreements. Consequently, Service Providers must be capable of detecting anomalies in their services in a timely fashion, while accommodating for scale. Around 10 thousand L3 VPNs in our Swisscom use case. Long-lasting outages, detected by the customer before the service provider, are detrimental to the perception of service quality, and may dramatically impact the customer business.

The goal of the presented architecture is to provide an anomaly detection solution that scales while being flexible on the following aspects: (i) the dimensions that must be used to detect anomalies are multiple; (ii) VPN customers wear different profiles in terms of normal and abnormal values for such dimensions; (iii) the amount of information collected to produce values for such dimensions is extremely large in such deployments: around 175 thousand messages/second in our use case; (iv) the operating costs for managing an anomaly detection solution must be kept low; and (v) the networking platforms providing the service may come from different vendors and have different monitoring capabilities.

The remainder paper is structured as follows. In section 2, we define what is considered a network anomaly and present the associated challenges behind its detection. In Section 3, we describe the Daisy architecture. In Section 4, we review the ongoing IETF efforts aimed at filling the gaps for a fully open, standard, Anomaly Detection (AD) implementation. And finally, in section 5, we present the first results of Daisy deployment at Swisscom.

2 PROBLEM STATEMENT

We describe some of the challenges associated with customer diversity, and a non-exhaustive list of anomalies targeted by the base recipes from our limited proof of concept deployment setup.

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