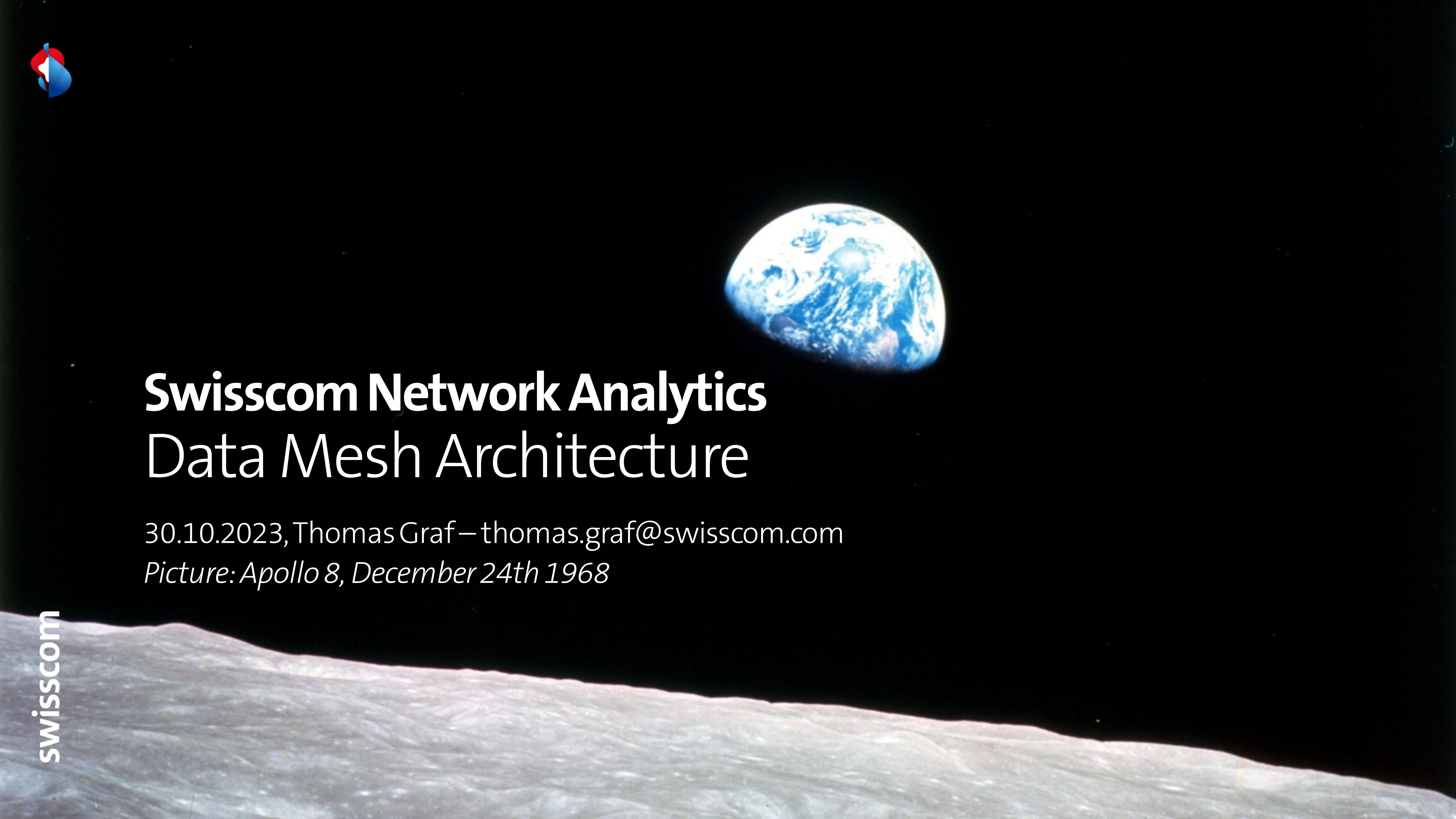


Swisscom Network Analytics

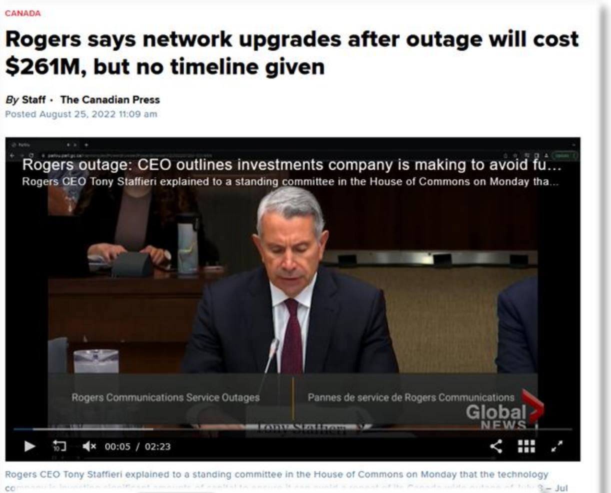
Data Mesh Architecture

30.10.2023, Thomas Graf – thomas.graf@swisscom.com

Picture: Apollo 8, December 24th 1968

Nationwide Network Outages everywhere

Increasing in impact and duration - hinting Network Visibility deficiencies



CANADA

Rogers says network upgrades after outage will cost $261M, but no timeline given

By Staff · The Canadian Press

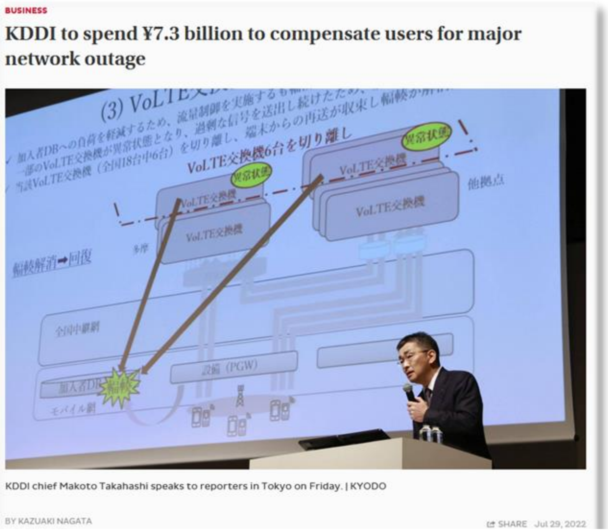
Posted August 25, 2022 11:09 am

Rogers outage: CEO outlines investments company is making to avoid fu...

Rogers CEO Tony Staffieri explained to a standing committee in the House of Commons on Monday tha...

Rogers Communications Service Outages
Pannes de service de Rogers Communications
Global NEWS
00:05 / 02:23

Rogers CEO Tony Staffieri explained to a standing committee in the House of Commons on Monday that the technology



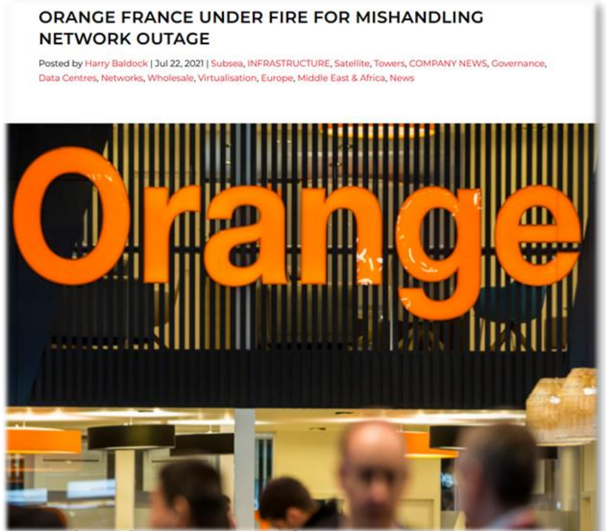
BUSINESS

KDDI to spend ¥7.3 billion to compensate users for major network outage

KDDI chief Makoto Takahashi speaks to reporters in Tokyo on Friday. | KYODO

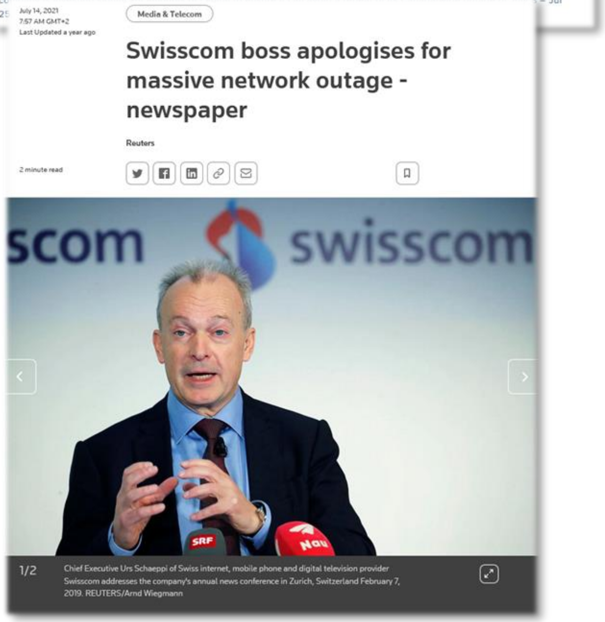
BY KAZUAKI NAGATA

SHARE Jul 29, 2022



ORANGE FRANCE UNDER FIRE FOR MISHANDLING NETWORK OUTAGE

Posted by Harry Baldock | Jul 22, 2021 | Subsea, INFRASTRUCTURE, Satellite, Towers, COMPANY NEWS, Governance, Data Centres, Networks, Wholesale, Virtualisation, Europe, Middle East & Africa, News



July 14, 2021
7:57 AM GMT+2
Last Updated a year ago

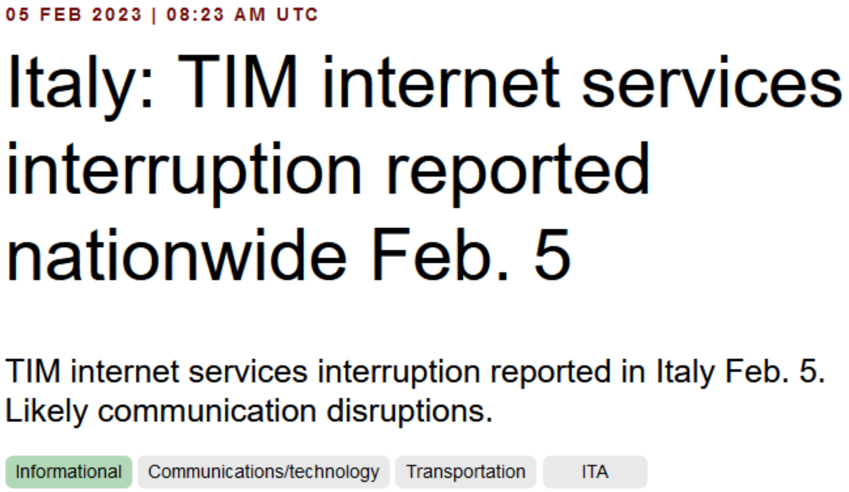
Media & Telecom

Swisscom boss apologises for massive network outage - newspaper

Reuters

2 minute read

1/2 Chief Executive Urs Schaeppi of Swiss internet, mobile phone and digital television provider Swisscom addresses the company's annual news conference in Zurich, Switzerland February 7, 2019. REUTERS/Arnd Wiegmann

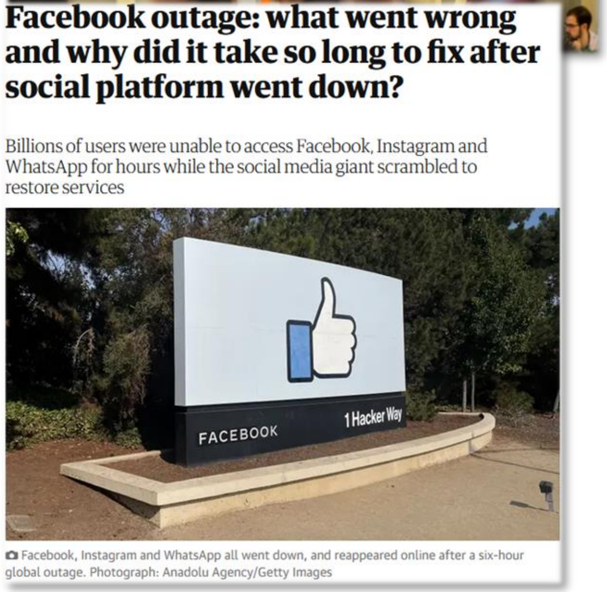


05 FEB 2023 | 08:23 AM UTC

Italy: TIM internet services interruption reported nationwide Feb. 5

TIM internet services interruption reported in Italy Feb. 5. Likely communication disruptions.

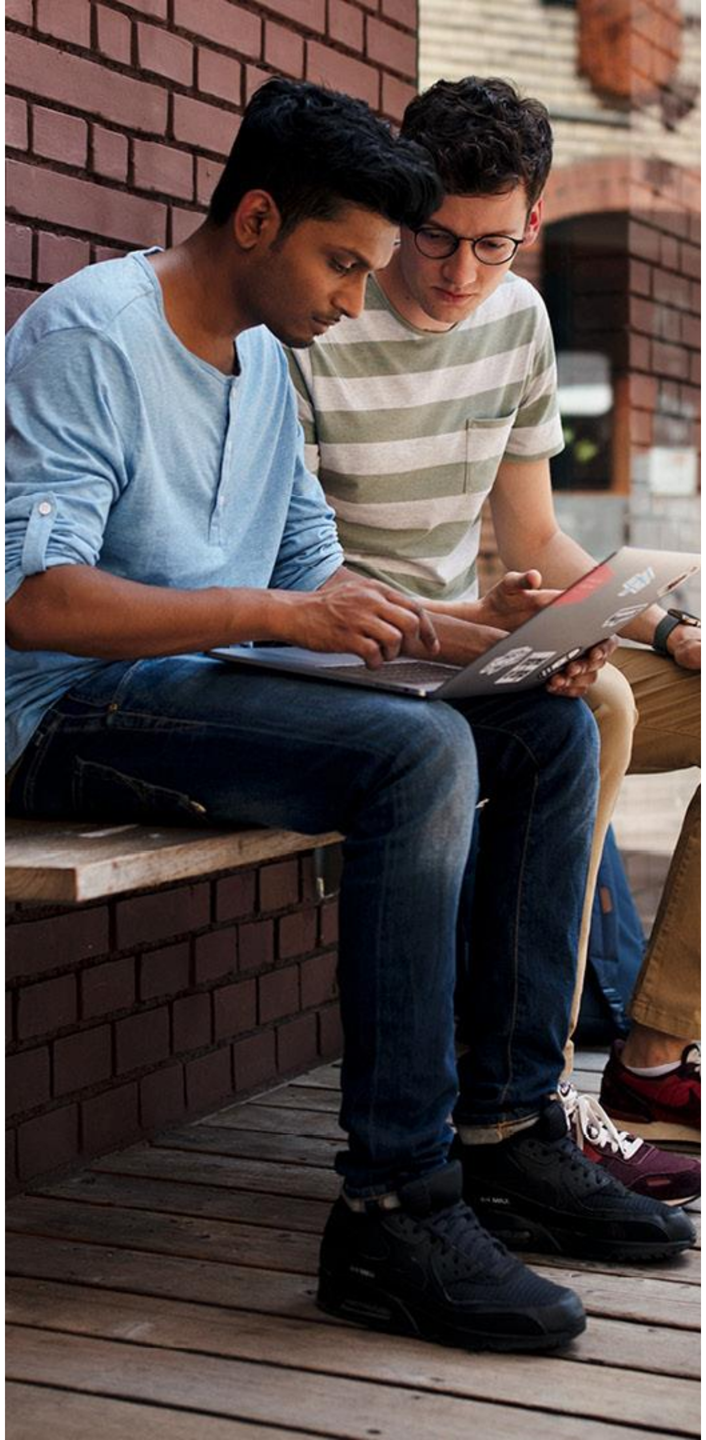
Informational | Communications/technology | Transportation | ITA



Facebook outage: what went wrong and why did it take so long to fix after social platform went down?

Billions of users were unable to access Facebook, Instagram and WhatsApp for hours while the social media giant scrambled to restore services

Facebook, Instagram and WhatsApp all went down, and reappeared online after a six-hour global outage. Photograph: Anadolu Agency/Getty Images



« The customer knows before Swisscom that there is service interruption.

Unable to recognize impact and root cause when configurational or operational network changes occur.

Swisscom suffers reputation damage.
We need to work together to mediate. »

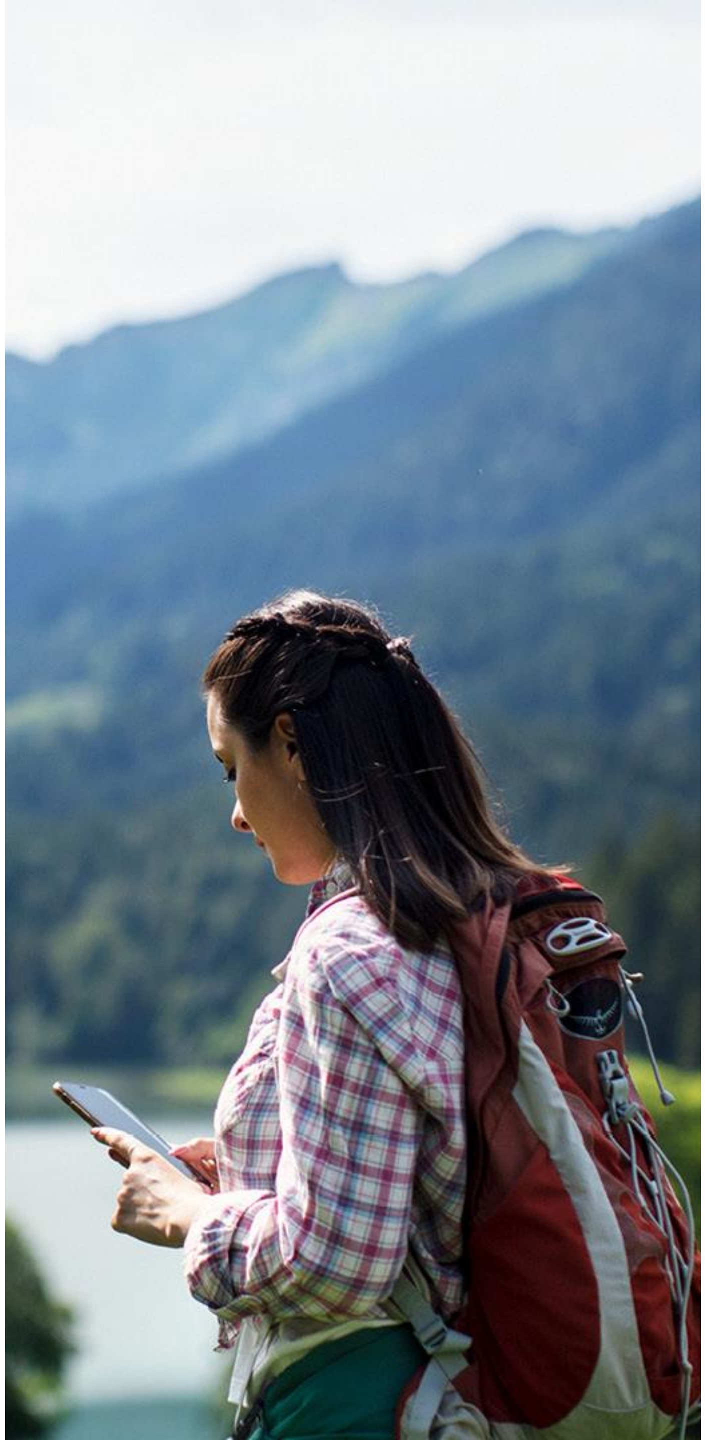
Markus Reber
Head of Networks at Swisscom

« At IETF only 9.85% of the activities are related to network automation and monitoring.

We are still using protocols designed 40 years ago to manage networks.

IP network protocols are not made to expose metrics for analytics. IPFIX and BGP monitoring protocol are the rare exception. »

Thomas Graf
Distinguished Network Engineer
and Network Analytics Architect at Swisscom



“ It is our duty to recognize service interruption before our customer does.

Why do we still often fail to be first ? “

Network Analytics Use Cases

What they are and how they relate

Network Data Collection

- Enables analytical use cases

Verification, Troubleshooting and Notification

- Dashboard, query and drill down on operational metrics

Network Anomaly Detection

- State change for connectivity services

Network Service Level Indicator and Objective

- State and state objective for connectivity services

Network Visualization

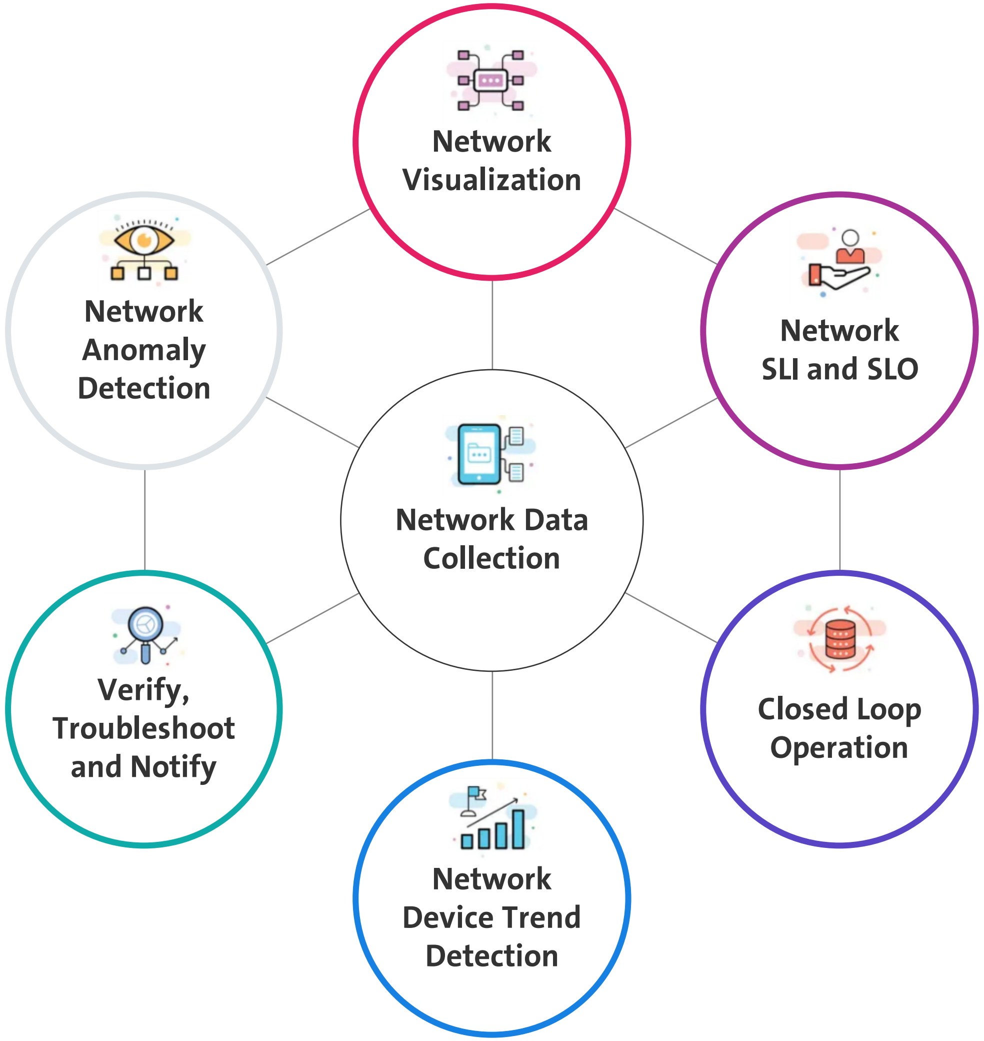
- Eases overview and access of metrics to humans

Network Device Trend Detection

- Tracks and predicts critical devices resources

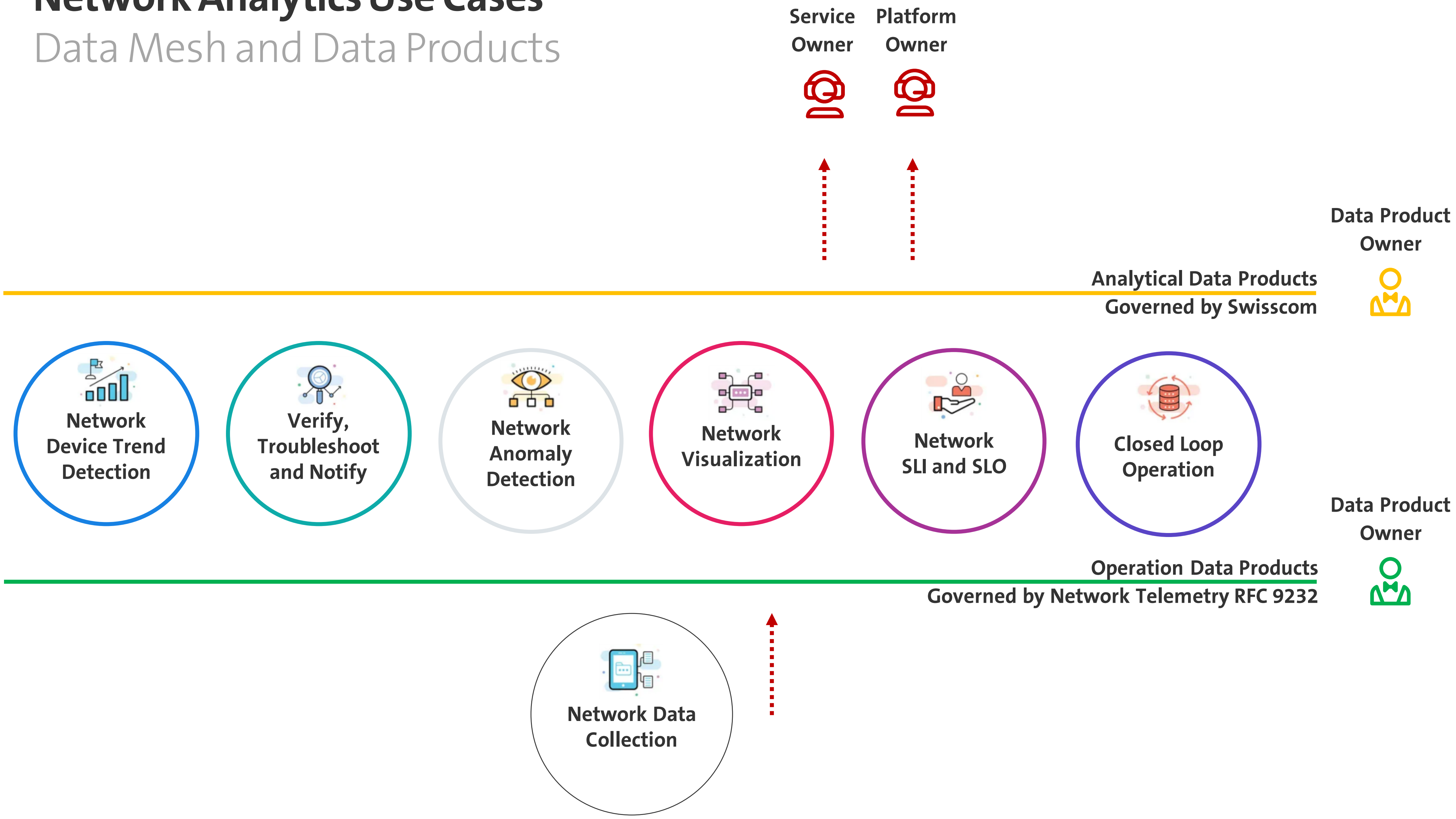
Closed Loop Operation

- Automates network verification


Network Visualization
Network Anomaly Detection
Network SLI and SLO
Network Data Collection
Verify, Troubleshoot and Notify
Closed Loop Operation
Network Device Trend Detection

Network Analytics Use Cases

Data Mesh and Data Products



Service Owner

Platform Owner

Data Product Owner

Analytical Data Products
Governed by Swisscom

Network Device Trend Detection

Verify, Troubleshoot and Notify

Network Anomaly Detection

Network Visualization

Network SLI and SLO

Closed Loop Operation

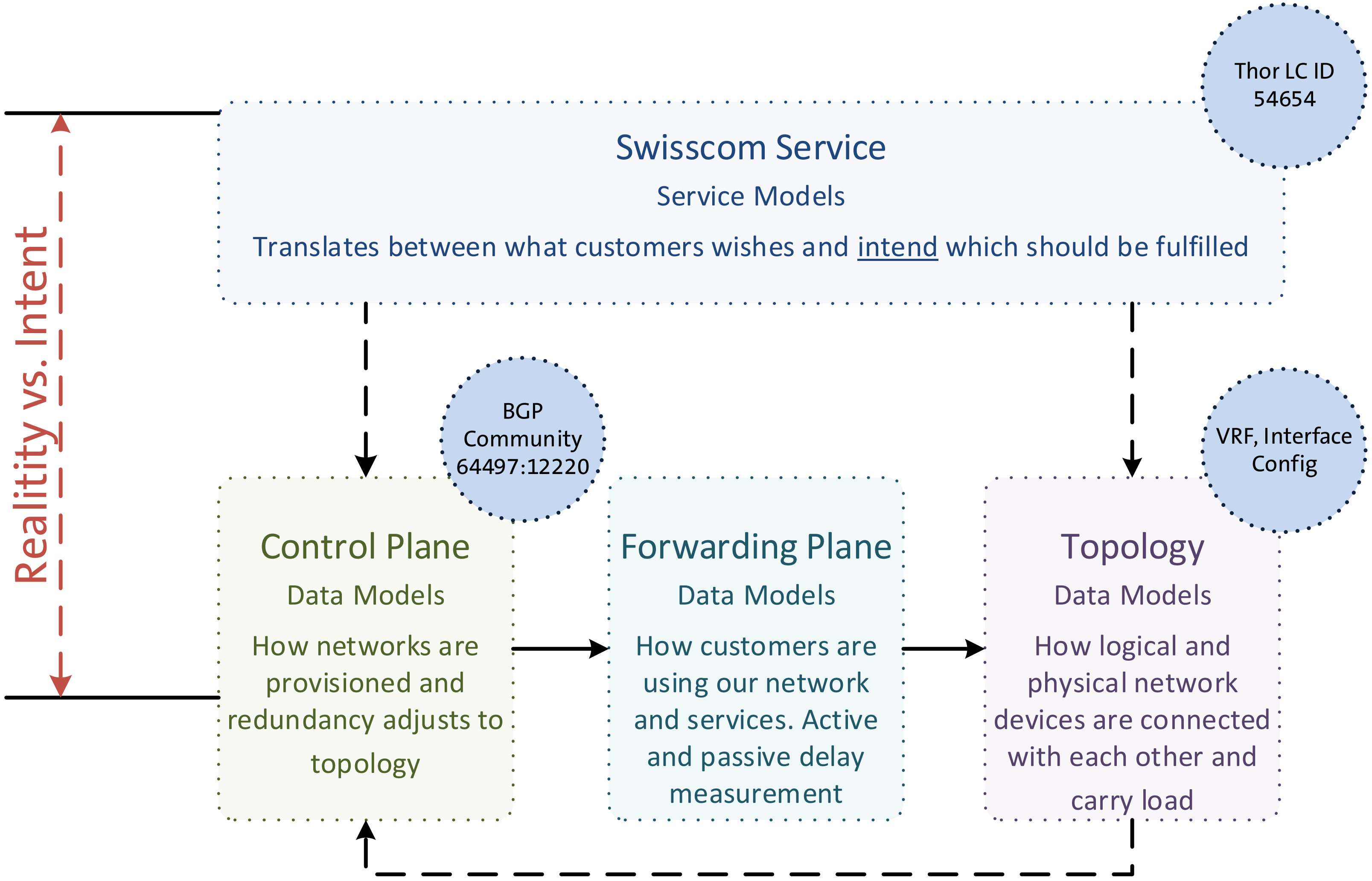
Data Product Owner

Operation Data Products
Governed by Network Telemetry RFC 9232

Network Data Collection

Data Collection with Network Telemetry

Structured metrics enable informed decision-making



Realitity vs. Intent

Thor LC ID 54654

Swisscom Service

Service Models

Translates between what customers wishes and intend which should be fulfilled

BGP Community 64497:12220

VRF, Interface Config

Control Plane

Data Models

How networks are provisioned and redundancy adjusts to topology

Forwarding Plane

Data Models

How customers are using our network and services. Active and passive delay measurement

Topology

Data Models

How logical and physical network devices are connected with each other and carry load

Network Telemetry:

- A data collection framework where the network device pushes its metrics to Big Data. Defined in RFC 9232.

Data Modelling:

- Key for Big Data correlation to understand and react in the right context
  - Are interface drops bad?
  - How should we react?

“ The solution comes with innovators.

That's why Swisscom cooperates at IETF with network operators, vendors and universities. “

Collaboration for tomorrows Network Analytics

Swisscom
Network Operator

INSA Lyon
University

Bell Canada
Network Operator

Huawei
Network Vendor

ETH Zürich
University

Cisco
Network Vendor

NTT
Network Operator

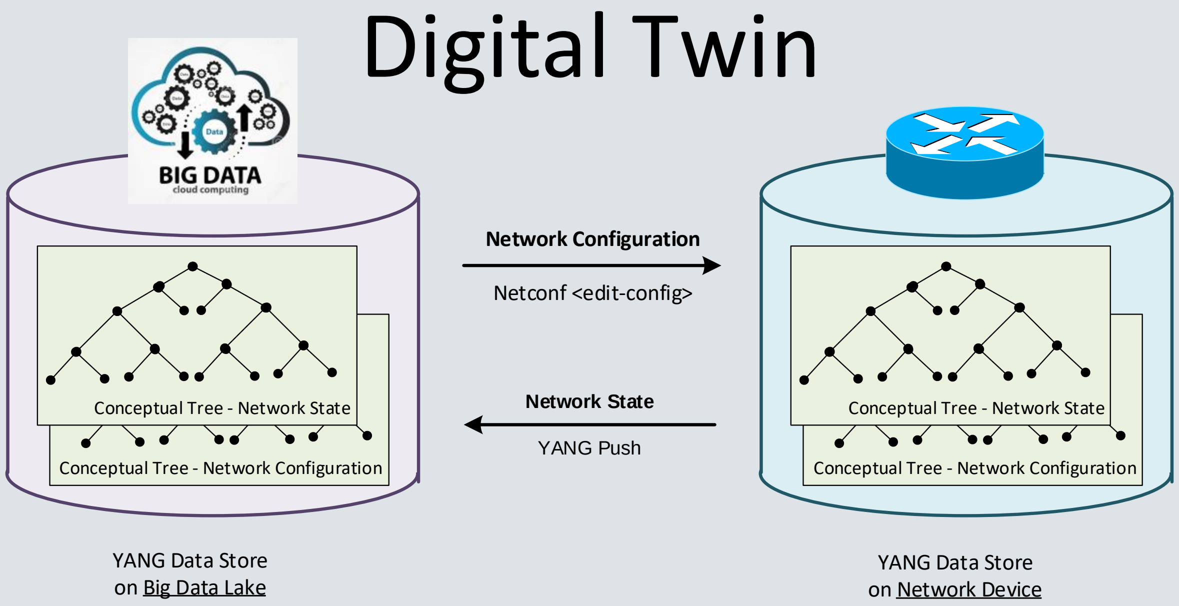
Imply
Imply Druid

Confluent
Apache Kafka

Clickhouse
OLAP DBMS

YANG Datastores enables Closed Loop Operation

Automated data correlation – what else?



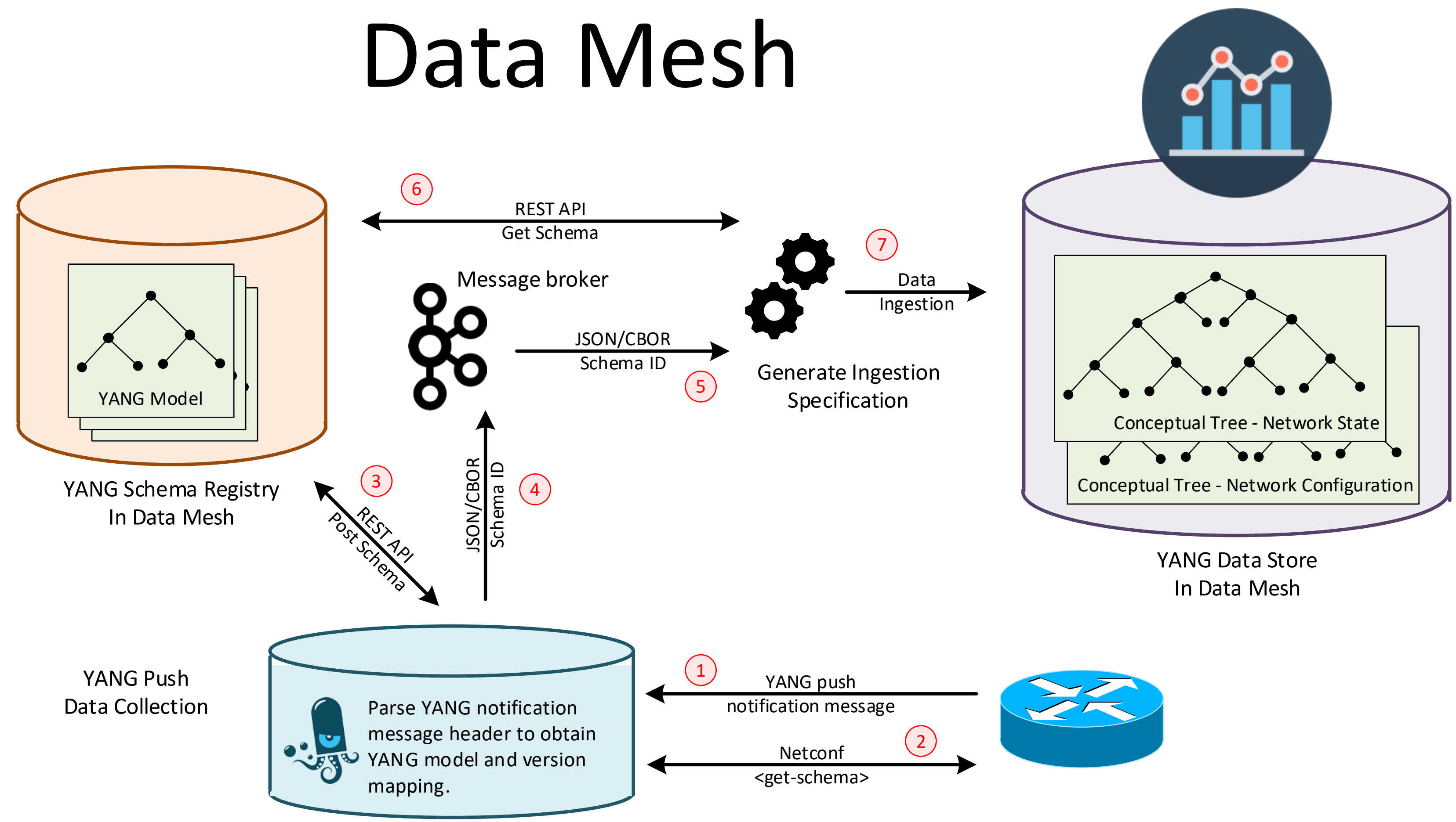
YANG is a data modelling language which will not only transform how we managed our networks; it will transform also how we manage our services.

News: **20 industry leading colleagues** from 4 network operators, 2 network and 4 analytics providers, and 2 universities **commit on a project to integrate YANG and CBOR into data mesh**. **Next update IETF 118 Prague.**

Automated networks can only run with a common data model. A digital twin YANG data store enables a comparison between intend and reality. Schema preservation enables closed loop operation. Closed Loop is like an autopilot on an airplane. We need to understand what the flight envelope is to keep the airplane within. Without, we crash.

When Big Data and Network become one

Marrying two messaging protocols



Preserve YANG data module definition throughout the data processing chain.

Enable automated data correlation among management, forwarding and control-plane **for anomaly detection.**

Simplify YANG push network data collection at high scale with low impact. **Suited for nowadays distributed forwarding systems.**

Support of Hostname and Sequencing in YANG Notifications
draft-tgraf-netconf-notif-sequencing

Support of Network Observation Timestamping in YANG Notifications
draft-tgraf-yang-push-observation-time

Support of Versioning in YANG Notifications Subscription
draft-netconf-yang-notifications-versioning

UDP-based Transport for Configured Subscriptions
draft-ietf-netconf-udp-notif

Subscription to Distributed Notifications
draft-ietf-netconf-distributed-notif

Yang
Kafka

BMP Covering all RIB's

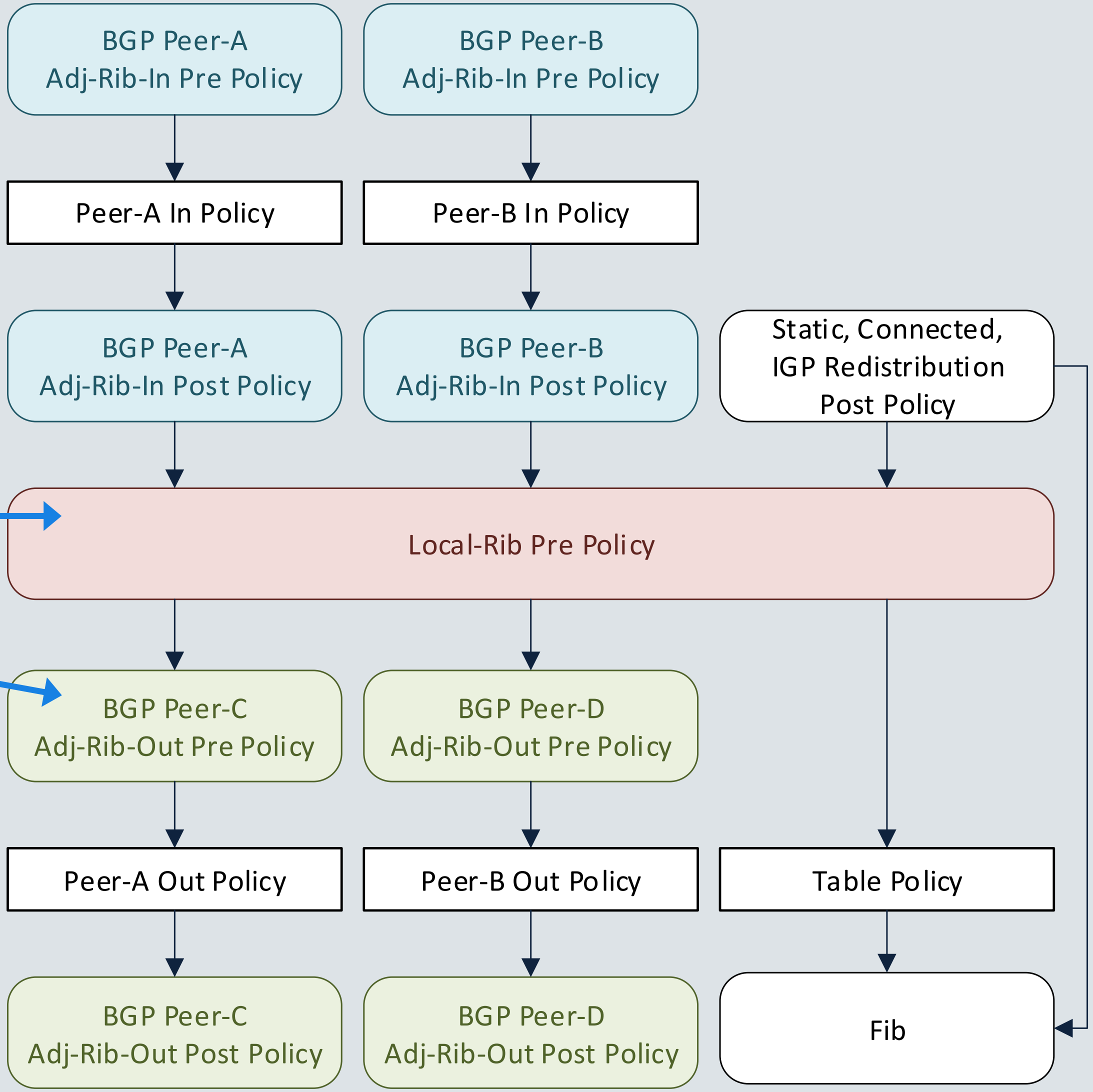
Extends much needed RIB coverage

BGP route exposure without BMP is a challenge of the first order:

- Only best path is exposed (missing best-external and ECMP routes)
- Next-hop attribute not preserved all the time
- Filtering between RIB's not visible

- **Support for Local RIB in BGP Monitoring Protocol** https://datatracker.ietf.org/doc/html/rfc9069
- **Support for Adj-RIB-Out in BGP Monitoring Protocol** https://tools.ietf.org/html/rfc8671

Adj-RIB-Out an RFC since November 2019. Local RIB since February 2022. Juniper, Huawei and Nokia have public releases available supporting both. Cisco for Local RIB since August 2023.



BGP Peer-A Adj-Rib-In Pre Policy
BGP Peer-B Adj-Rib-In Pre Policy
Peer-A In Policy
Peer-B In Policy
BGP Peer-A Adj-Rib-In Post Policy
BGP Peer-B Adj-Rib-In Post Policy
Static, Connected, IGP Redistribution Post Policy
Local-Rib Pre Policy
BGP Peer-C Adj-Rib-Out Pre Policy
BGP Peer-D Adj-Rib-Out Pre Policy
Peer-A Out Policy
Peer-B Out Policy
Table Policy
BGP Peer-C Adj-Rib-Out Post Policy
BGP Peer-D Adj-Rib-Out Post Policy
Fib

BMP with extended TLV support

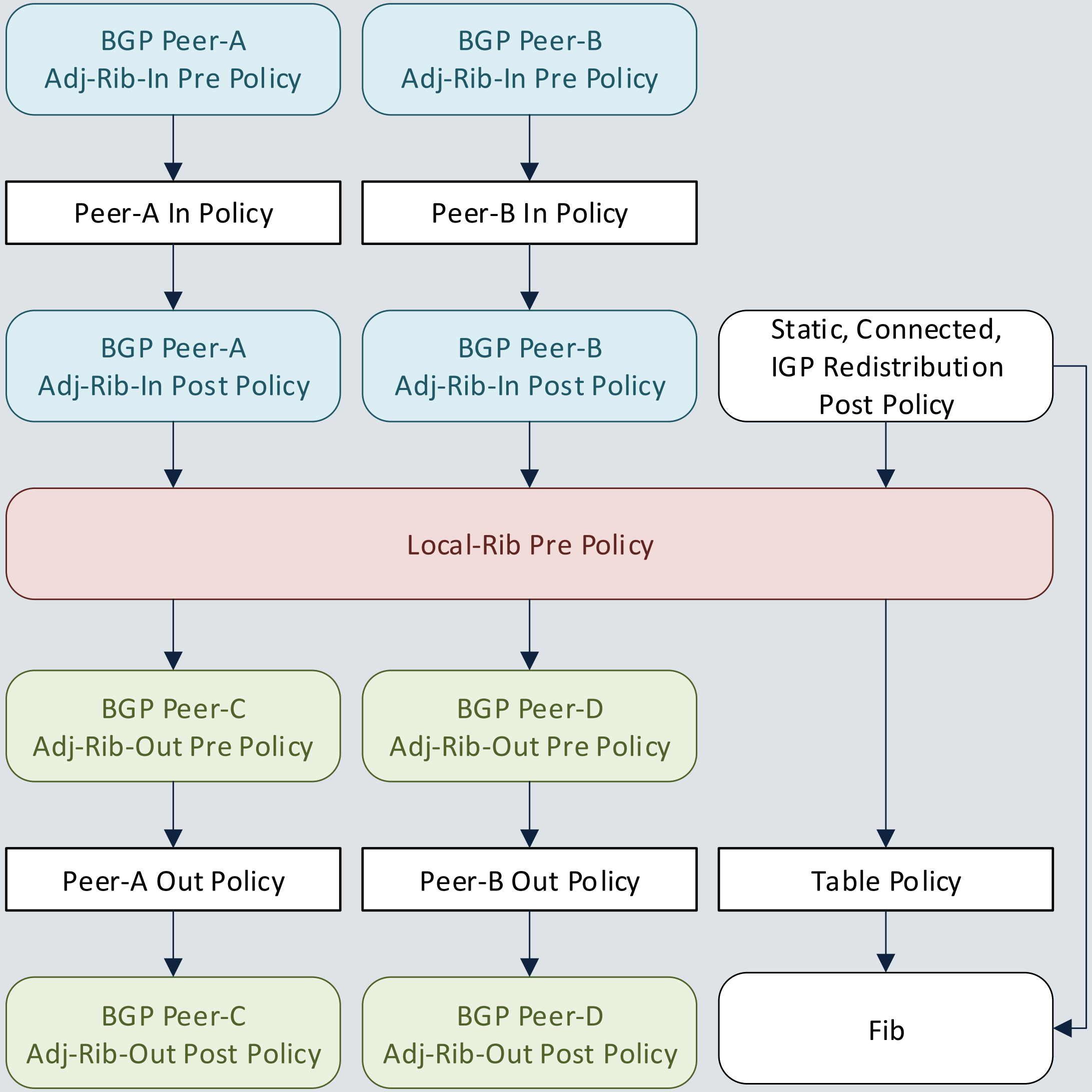
Brings visibility into FIB's and route-policies

Knowing all the routes in all the RIB's brings the new challenge

- That we don't know how they are being used in the FIB/RIB (which one is best, best-external, ECMP, backup)
- That we don't know which route-policy permitted/denied/changed which prefix/attribute

- **TLV support for BMP Route Monitoring and Peer Down Messages**
  https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-tlv-ebit
- **Support for Enterprise-specific TLVs in the BGP Monitoring Protocol**
  https://tools.ietf.org/html/draft-lucente-grow-bmp-tlv-ebit
- **BMP Extension for Path Marking TLV**
  https://datatracker.ietf.org/doc/html/draft-ietf-grow-bmp-path-marking-tlv

All documents are GROW working group documents. Huawei has test code available. Frrouting is about to merge code.


BGP Peer-A Adj-Rib-In Pre Policy
BGP Peer-B Adj-Rib-In Pre Policy
Peer-A In Policy
Peer-B In Policy
BGP Peer-A Adj-Rib-In Post Policy
BGP Peer-B Adj-Rib-In Post Policy
Static, Connected, IGP Redistribution Post Policy
Local-Rib Pre Policy
BGP Peer-C Adj-Rib-Out Pre Policy
BGP Peer-D Adj-Rib-Out Pre Policy
Peer-A Out Policy
Peer-B Out Policy
Table Policy
BGP Peer-C Adj-Rib-Out Post Policy
BGP Peer-D Adj-Rib-Out Post Policy
Fib

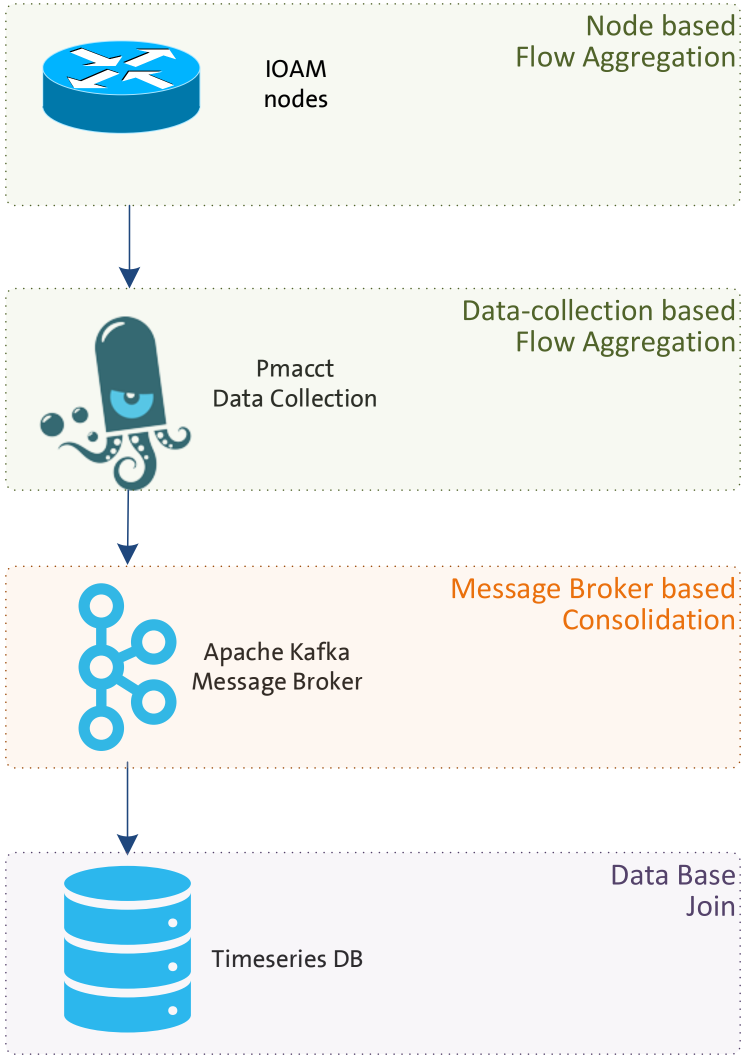
IPFIX Covering Segment Routing

For MPLS-SR, SRv6 and On-path Delay

SRv6 is commonly standardized, network vendors implementations are available and network operators are at various stages in their deployments, missing data-plane visibility though.

Segment Routing coverage in IPFIX brings visibility for:

- Which routing protocol provided the label or IPv6 Segment in the SR domain.
- The active Segment where the packet is forwarded to in the SRv6 Domain.
- The Segment List where the packet is going to be forwarded throughout the SRv6 Domain.
- The Endpoint Behavior describing how the packet is being forwarded in the SRv6 Domain.
- The Min, Max and Average On-path delay at each hop in the SR domain.


Node based Flow Aggregation
IOAM nodes
Data-collection based Flow Aggregation
Pmacct Data Collection
Message Broker based Consolidation
Apache Kafka Message Broker
Data Base Join
Timeseries DB

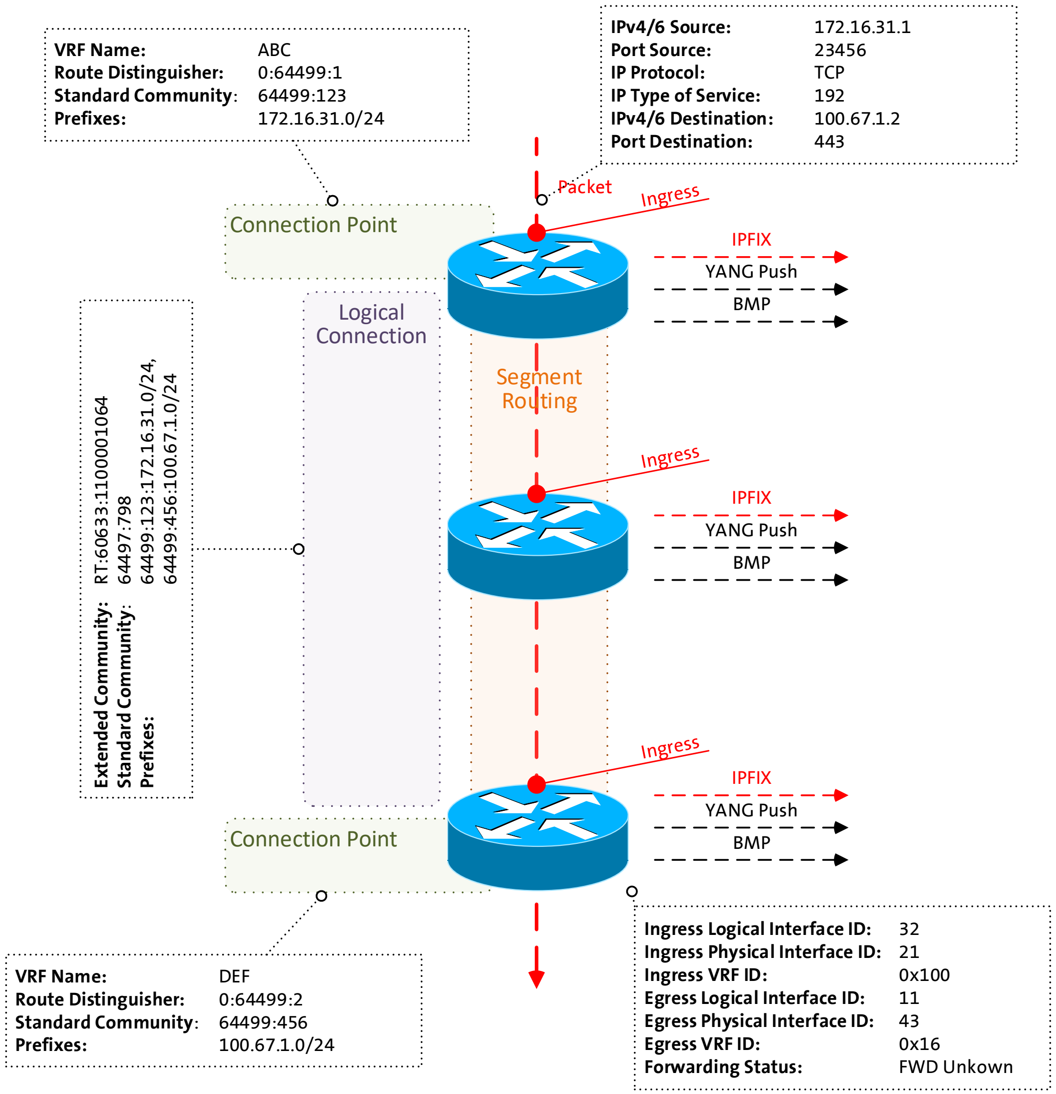
Export of MPLS Segment Routing Label Type Information in IPFIX
https://datatracker.ietf.org/doc/html/rfc9160

Export of Segment Routing IPv6 Information in IPFIX
https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-ipfix-srv6-srh

Export of Forwarding Path Delay in IPFIX
https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-ipfix-on-path-telemetry

Inband Telemetry with IPFIX Flow-Aggregation

Aggregate and sample as early as possible – Chose your Cardinality



VRF Name:	ABC
Route Distinguisher:	0:64499:1
Standard Community:	64499:123
Prefixes:	172.16.31.0/24

IPv4/6 Source:	172.16.31.1
Port Source:	23456
IP Protocol:	TCP
IP Type of Service:	192
IPv4/6 Destination:	100.67.1.2
Port Destination:	443

Extended Community:	RT:60633:1100001064
Standard Community:	64497:798
Prefixes:	64499:123:172.16.31.0/24, 64499:456:100.67.1.0/24

VRF Name:	DEF
Route Distinguisher:	0:64499:2
Standard Community:	64499:456
Prefixes:	100.67.1.0/24

Ingress Logical Interface ID:	32
Ingress Physical Interface ID:	21
Ingress VRF ID:	0x100
Egress Logical Interface ID:	11
Egress Physical Interface ID:	43
Egress VRF ID:	0x16
Forwarding Status:	FWD Unkown

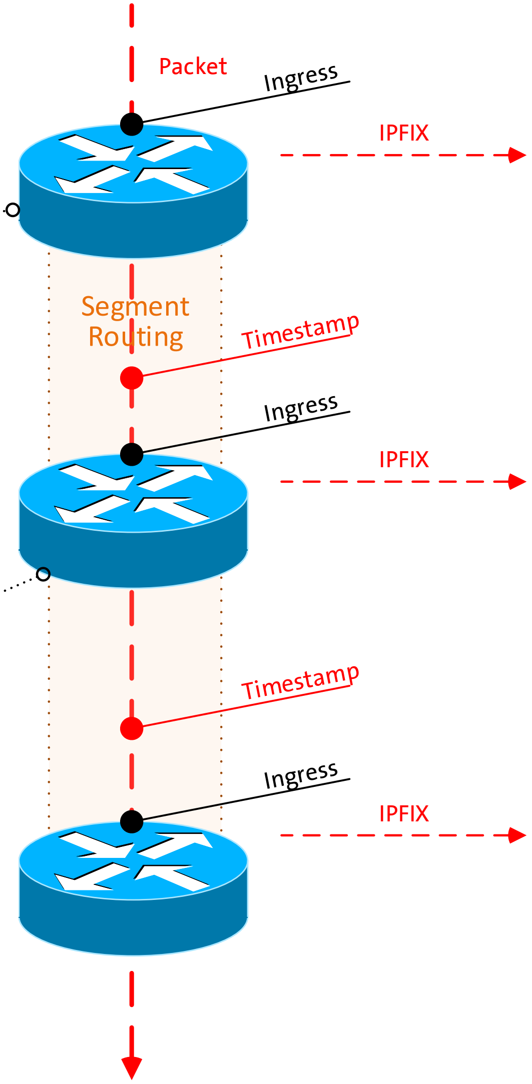
- IPFIX defines two key data engineering tools to reduce collected and exported amount of data. **Sampling and Aggregation.** Enabling **a statistical view from the network usage.** Also called **connectivity matrix.**
- IPFIX **measures packets and bytes** and give **device and control-plane context**.
- **With Inband Telemetry**, iOAM, Path Tracing and iFIT, **delay can be measured** actively (probing) or passively. Metrics are exposed on every node, postcards or only at the last node (passport).
- **IPFIX lacks the ability to export delay.** A key element for monitoring Customer Service Level Agreements.
- **Inband Telemetry lacks Flow Aggregation support** as defined in RFC 7015. Therefore, **scalability** in terms of data export and collection is **drastically limited** today.
- draft-tgraf-opsawg-ipfix-inband-telemetry enables IPFIX to export delay while preserving the ability to aggregate and also **adds the Inband Telemetry path delay metric definition** in the performance registry for proper delay definition.

Measure delay and give network context

Enabling a statistical network delay view

IPv4/6 Source:	172.16.31.1
Port Source:	23456
IP Protocol:	TCP
IP Type of Service:	192
IPv4/6 Destination:	100.67.1.2
Port Destination:	443
Ingress Logical Interface ID:	32
Ingress Physical Interface ID:	21
Ingress VRF ID:	0x100
Egress Logical Interface ID:	11
Egress Physical Interface ID:	43
Egress VRF ID:	0x16
Forwarding Status:	FWD Unkown

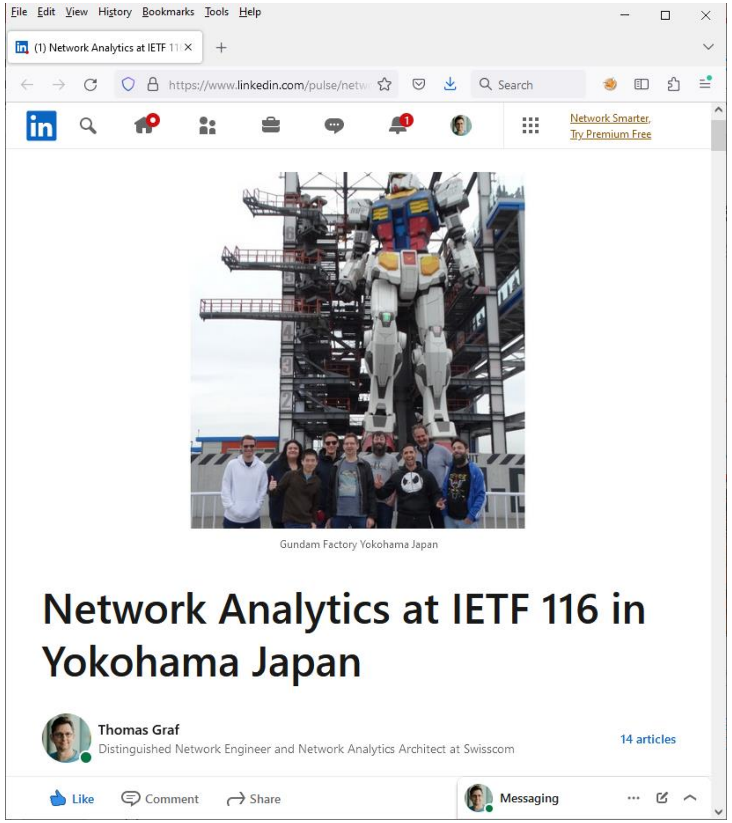
IPv4/6 Source:	172.16.31.1
Port Source:	23456
IP Protocol:	TCP
IP Type of Service:	192
IPv4/6 Destination:	100.67.1.2
Port Destination:	443
Ingress Logical Interface ID:	32
Ingress Physical Interface ID:	21
Ingress VRF ID:	0x16
Egress Logical Interface ID:	11
Egress Physical Interface ID:	43
Egress VRF ID:	0x16
Forwarding Status:	FWD Unkown
SID List:	17001, 34002
Delay Min	1
Delay Sum	5
Delay Max	7



- Packets are captured ingress with an optional sampler, data-plane dimensions extracted, enriched with device and control-plane dimensions and **added with a unique flow ID to a flow cache on the node for aggregation.**
- The data-plane dimensions answers **which packet**. The control-plane **which service**. The device dimensions **where in the network**.
- In case of Inband Telemetry, **a timestamp and optionally a direct export tag is added** to the packet header when entering the Inband Telemetry domain.
- Each subsequent packet for the same flow increases byte and packet count. Each new flow creates a new flow ID in the flow cache.
- In case of Inband Telemetry, At each node in transit (postcard) or only at the last node (passport), **the delay is calculated by comparing the timestamp in the packet and when packet is received** on the node**. Delay is populated into the flow cache besides packet and byte count.**

IETF 116/117– Network Analytics Development

SRv6 Data Plane Visibility and YANG/Kafka Integration


File Edit View History Bookmarks Tools Help
(1) Network Analytics at IETF 11
https://www.linkedin.com/pulse/netw
Search
Network Smarter, Try Premium Free
Like
Comment
Share
Messaging

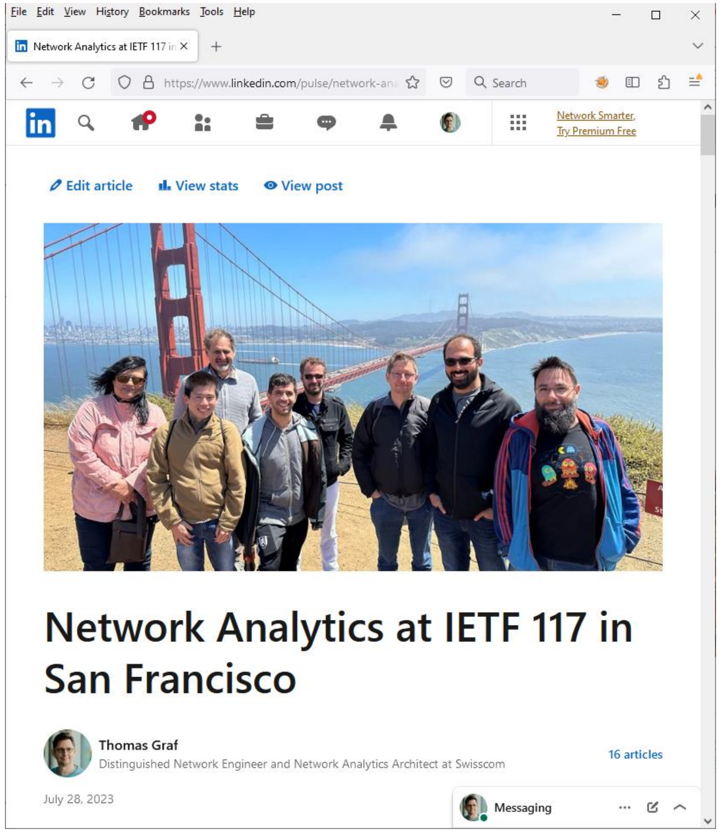
Gundam Factory Yokohama Japan

Network Analytics at IETF 116 in Yokohama Japan

Thomas Graf
Distinguished Network Engineer and Network Analytics Architect at Swisscom

14 articles

https://www.linkedin.com/pulse/network-analytics-ietf-116-yokohama-thomas-graf/


File Edit View History Bookmarks Tools Help
Network Analytics at IETF 117 in
https://www.linkedin.com/pulse/network-an
Search
Network Smarter, Try Premium Free
Edit article
View stats
View post
Messaging

Network Analytics at IETF 117 in San Francisco

Thomas Graf
Distinguished Network Engineer and Network Analytics Architect at Swisscom

16 articles

July 28, 2023

https://www.linkedin.com/pulse/network-analytics-ietf-117-san-francisco-thomas-graf/

- **5x BMP drafts and 1 RFC** at GROW working group. Bringing RIB and route-policy dimensions into BMP and increase scale.
- **6x YANG push drafts** at NETCONF working group.
- **2x IPFIX Segment Routing On-path delay draft and 1 RFC** at OPSAWG working group.
- **2x IOAM DEX** drafts at IPPM working group.
- **Network Anomaly Detection** code development.
- ✓ **YANG push udp-notif, BMP, IPFIX SRv6, On-Path and IOAM** open-source running code.