Swisscom: Network Incident Network Analytics Postmortem

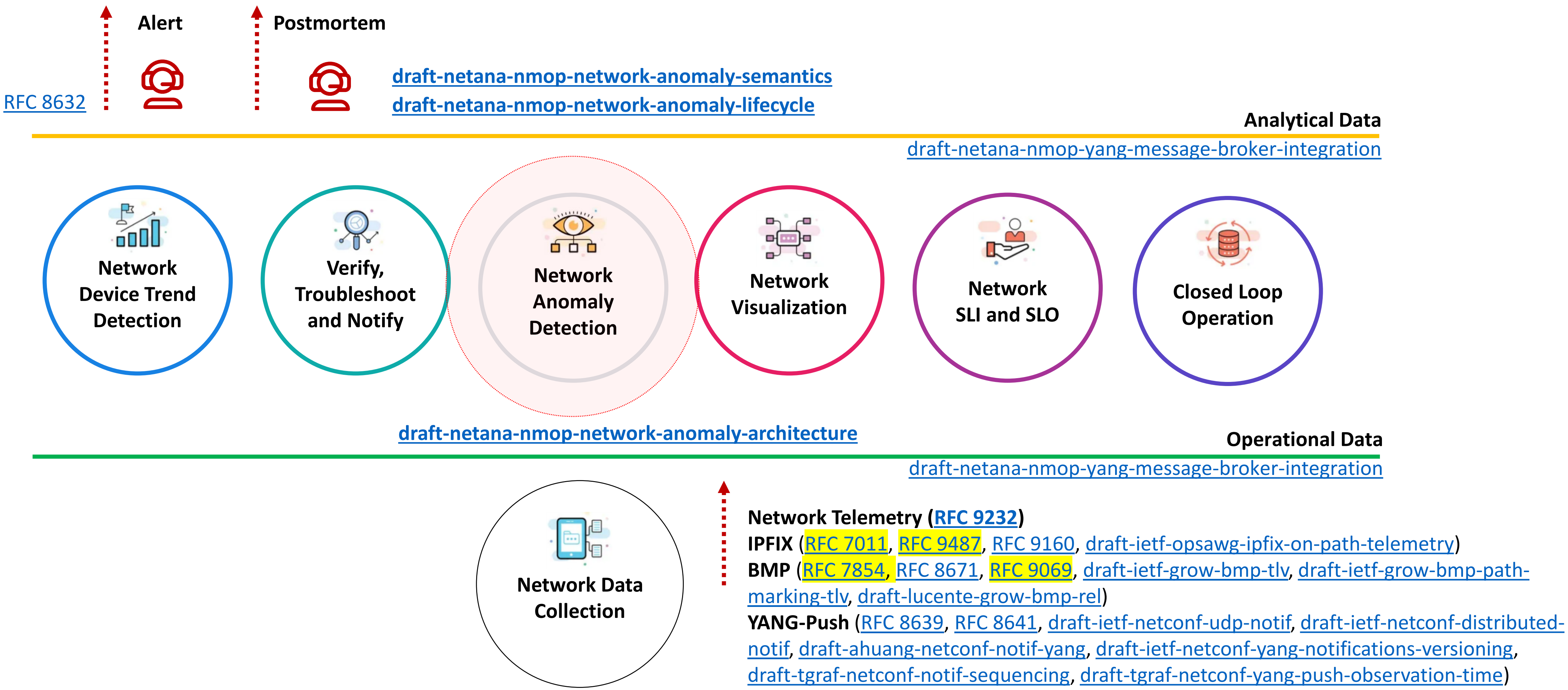
Describes an incident in terms of
what happened,
which operational metrics where available,
which analytical metrics described the symptoms and
what improvements in the network anomaly detection system and network telemetry protocols are proposed.

thomas.graf@swisscom.com

02. November 2024

Data Mesh organizes Data in Organizations

Enables Network Analytics use cases



Alert

Postmortem

RFC 8632

draft-netana-nmop-network-anomaly-semantics

draft-netana-nmop-network-anomaly-lifecycle

Analytical Data

draft-netana-nmop-yang-message-broker-integration

Network Device Trend Detection

Verify, Troubleshoot and Notify

Network Anomaly Detection

Network Visualization

Network SLI and SLO

Closed Loop Operation

draft-netana-nmop-network-anomaly-architecture

Operational Data

draft-netana-nmop-yang-message-broker-integration

Network Data Collection

Network Telemetry (RFC 9232)

IPFIX (RFC 7011, RFC 9487, RFC 9160, draft-ietf-opsawg-ipfix-on-path-telemetry)

BMP (RFC 7854, RFC 8671, RFC 9069, draft-ietf-grow-bmp-tlv, draft-ietf-grow-bmp-path-marking-tlv, draft-lucente-grow-bmp-rel)

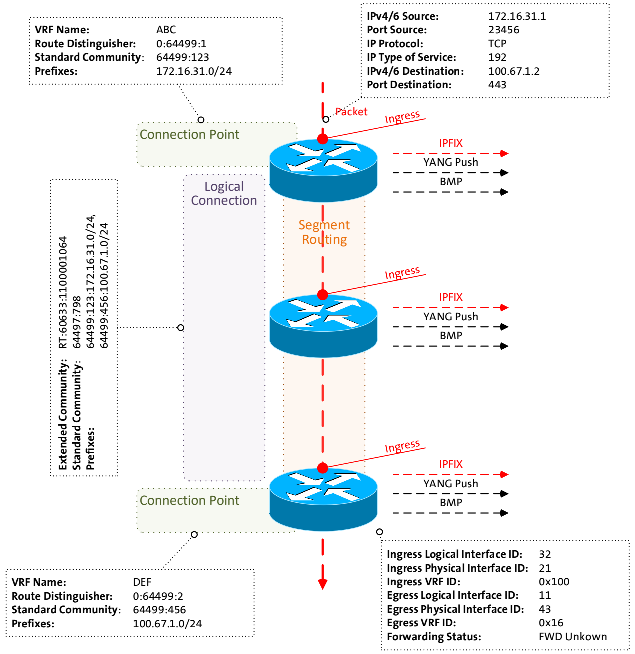
YANG-Push (RFC 8639, RFC 8641, draft-ietf-netconf-udp-notif, draft-ietf-netconf-distributed-notif, draft-ahuang-netconf-notif-yang, draft-ietf-netconf-yang-notifications-versioning, draft-tgraf-netconf-notif-sequencing, draft-tgraf-netconf-yang-push-observation-time)

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 »



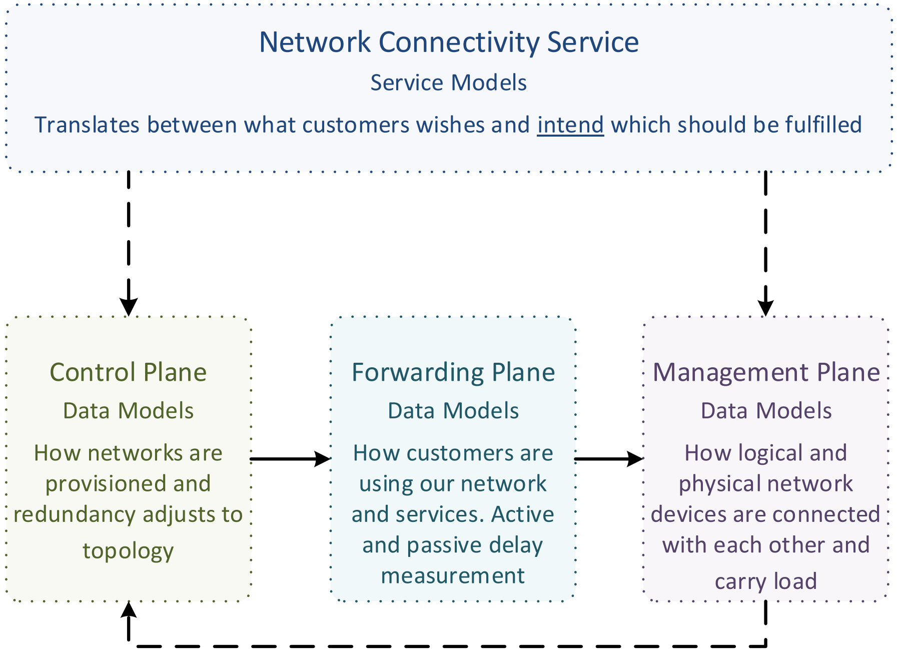
VRF Name:	ABC
Route Distinguisher:	0:64499:1
Standard Community:	64499:123
Prefixes:	172.16.31.0/24

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



Network Connectivity Service

Service Models

Translates between what customers wishes and intend which should be fulfilled

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

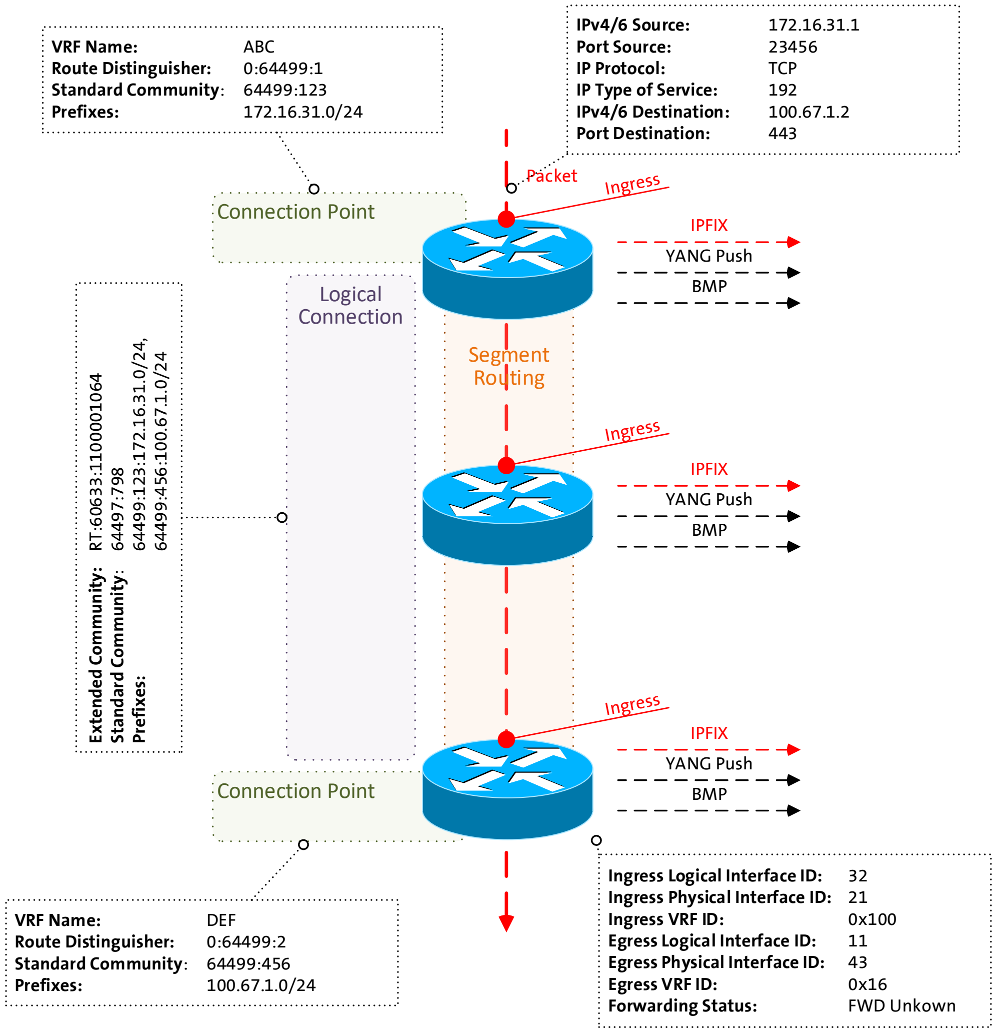
Management Plane

Data Models

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

Monitoring L3 VPN's with IPFIX, BMP and YANG Push

From Connectivity Service to Realtime Network Analytics



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

- **Connectivity Service perspective,** Connection Points are connected through Logical Connections.
- **From a BGP control-plane perspective,** IPv4/6 unicast prefixes in VRF's are tagged with BGP standard communities.
  - One BGP standard community to identify the Logical Connection. One BGP standard community to identify each Connection Point.
  - When IPv4/6 prefixes are exported from VRF's, a BGP route-distinguisher, BGP extended community route-targets and a SRv6 VPN SID for the IPv6 next-hop are allocated.
- **From a forwarding plane perspective,** when IPv4/6 unicast traffic is received from the edge at the SRv6 PE, a lookup is performed, the SRv6 VPN SID is obtained and IPv6 next-hop is added when forwarded to the core.
- **Swisscom collects** MPLS and SRv6 provider data plane, IPv4/6 unicast customer data-plane in IPFIX and at provider edge BGP VPNv4/6 unicast **in production** to perform real-time data correlation.

Problem Statement and Motivation

How it is being addressed in which document

Network Anomaly Detection

When operational or configurational changes in connectivity services are happening, the objective is to detect interruption at network operation faster than the users using those connectivity services

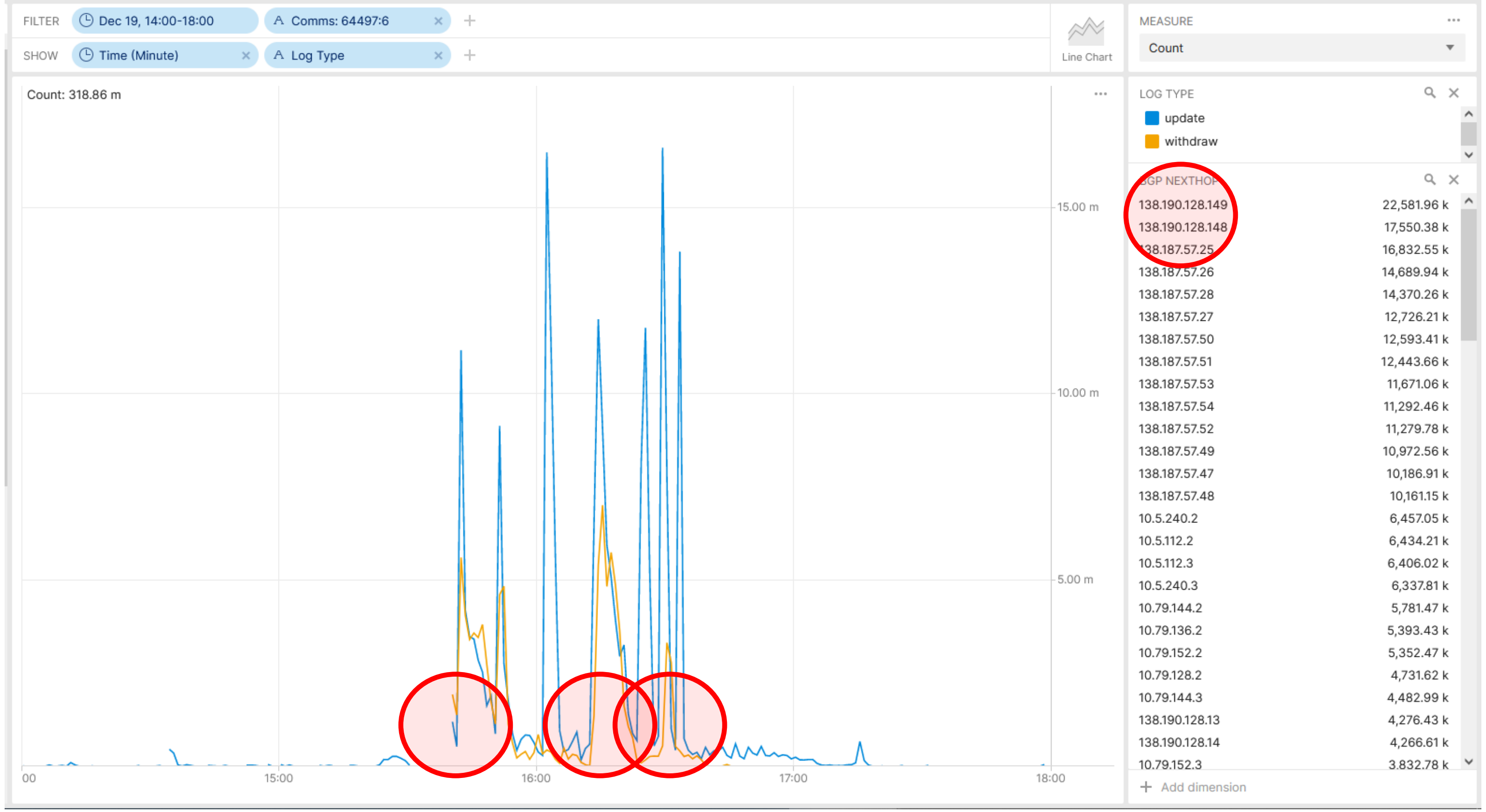
In order to achieve this objective, automation in network monitoring is required. This automation needs to monitor network changes holistically by monitoring all 3 network planes simultaneously and detect whether that change is service disruptive.

Through network incidents postmortems we network operators learn and improve so does network anomaly detection and supervised and semi-supervised machine learning. With more and more incidents the postmortem process demands automation and with the standardization of labeled network incident collaboration among network operators, vendors and academia is facilitated.

- ➢ draft-ietf-nmop-network-anomaly-architecture describes the motivation and architecture and the relationship to other two documents.
- ➢ draft-netana-nmop-network-anomaly-semantics defines Symptom semantics to enable standardized data exchange to validate results with network engineers and improve supervised and semi-supervised machine learning systems.
- ➢ draft-netana-nmop-network-anomaly-lifecycle describes on managing the lifecycle process, in order to facilitate network engineers to interact with the network anomaly detection system to refine the detection abilities over time.

Maximum Prefix BGP Peer State Change

What have happened



FILTER: Dec 19, 14:00-18:00 | Comms: 64497:6

SHOW: Time (Minute) | Log Type

MEASURE: Count

LOG TYPE

- 🟦 update
- 🟧 withdraw

BGP NEXTHOP	
138.190.128.149	22,581.96 k
138.190.128.148	17,550.38 k
138.187.57.25	16,832.55 k
138.187.57.26	14,689.94 k
138.187.57.28	14,370.26 k
138.187.57.27	12,726.21 k
138.187.57.50	12,593.41 k
138.187.57.51	12,443.66 k
138.187.57.53	11,671.06 k
138.187.57.54	11,292.46 k
138.187.57.52	11,279.78 k
138.187.57.49	10,972.56 k
138.187.57.47	10,186.91 k
138.187.57.48	10,161.15 k
10.5.240.2	6,457.05 k
10.5.112.2	6,434.21 k
10.5.112.3	6,406.02 k
10.5.240.3	6,337.81 k
10.79.144.2	5,781.47 k
10.79.136.2	5,393.43 k
10.79.152.2	5,352.47 k
10.79.128.2	4,731.62 k
10.79.144.3	4,482.99 k
138.190.128.13	4,276.43 k
138.190.128.14	4,266.61 k
10.79.152.3	3,832.78 k

BMP route-monitoring update/withdrawals on 64497:6

SOS

Long time ago, both a set of Inter-AS Option A ASBR routers started to log and notify through SNMP traps warning messages that 20% of the configured BGP maximum-prefix limit has been crossed. This has been visualized in an NMS with severity yellow and not being observed.

SOS

At 15:40 the configured limit has been reached and both redundant peers were shutdown 4 times for 10 minutes each at the same time.

At 15:41 Network Anomaly Detection observed on L3 VPN 64497:6 a potential issue with a concern score of 0.26 and at 16:02 reached the alert level of 0.30 and was not observed by 7x24 NOC.

At 15:41-45 network operation center noticed Swiss wide connectivity interruption on application level. **Unable to identify based on network metrics, suspecting due to scope a specific set of ASBR's and notified responsible platform team.**

At 16:10 ASBR team reached out to MPLS core team. **At 16:20 BGP** maximum prefix limit of peering was increased and peering state resolved.

Maximum Prefix BGP Peer State Change

Network Telemetry Coverage

IPFIX configured on P and PE MPLS-SR nodes on MPLS and IPv4/6 VRF unicast enabled interfaces. Capturing L3 **IPv4/6 and L2 Ethernet overlay customer data plane** and underlay MPLS provider data plane metrics on MPLS enabled interfaces, and IPv4/6 and L2 Ethernet overlay customer data plane metrics on IPv4/6 VRF unicast enabled interfaces.

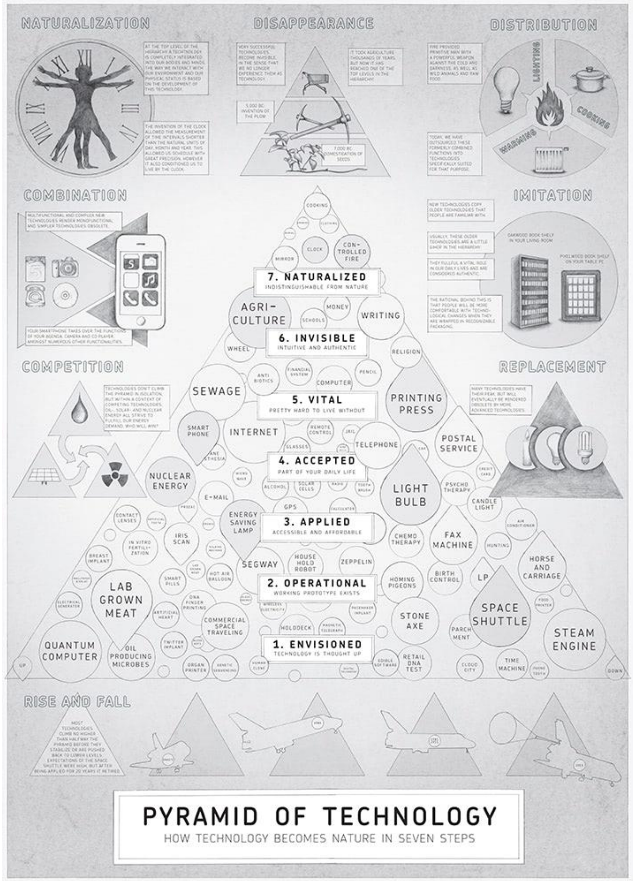
-> Shape, means that we are engaged in IETF standardization, vendor implementations and running code. IPv4/6 unicast customer data plane visibility is in vital, MPLS data plane visibility is in applied, On-Path delay is in operational stage.

BMP Adj-RIB In post-policy on BGP VPNv4 /6 and IPv4/6 VRF unicast peers and Local-RIB on all RIB's configured on MPLS PE's. BMP Adj-RIB In post-policy on BGP VPNv4 /6 peers on Route Reflectors configured.

-> Shape, means that we are engaged in IETF standardization, vendor implementations and running code. BMP Local RIB data plane visibility is in applied, BMP Path Marking is in operational stage.

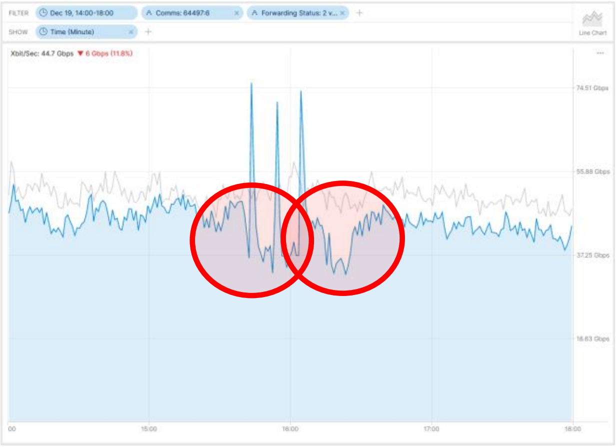
YANG Push Legacy on most nodes enabled but not relevant for this use case.

-> Take, means that current YANG-Push legacy implementation is used without any vendor code change and is in accepted stage. However, IETF YANG-Push is shape and is in operational state.

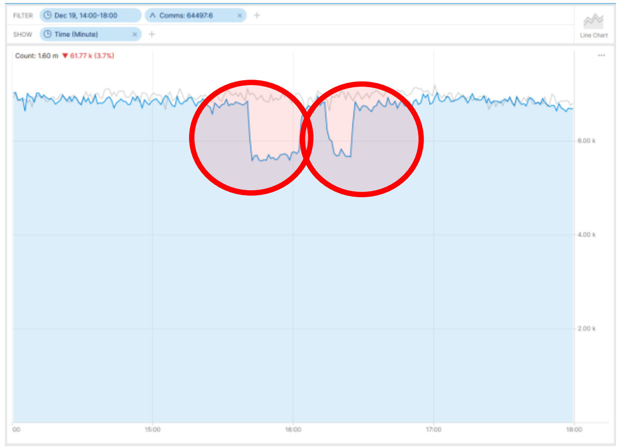


Postmortem, Maximum Prefix BGP Peer State Change

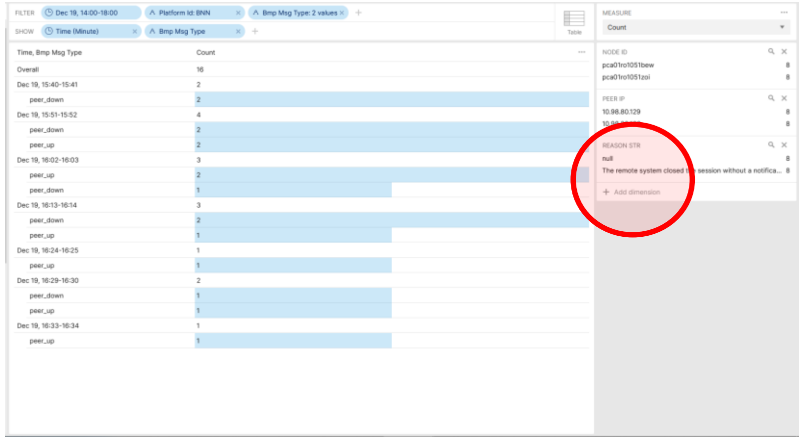
Which operational metrics covered



Missing Traffic 64497:6



Flow Count Drop 64497:6



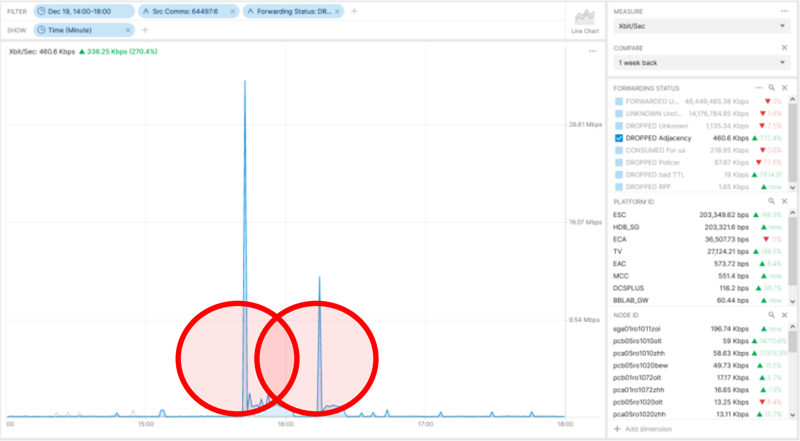
Time, Bmp Msg Type	Count
Overall	16
Dec 19, 15:40-15:41	2
peer_down	2
Dec 19, 15:51-15:52	4
peer_down	2
peer_up	2
Dec 19, 16:02-16:03	3
peer_up	2
peer_down	1
Dec 19, 16:13-16:14	3
peer_down	2
peer_up	1
Dec 19, 16:24-16:25	1
peer_up	1
Dec 19, 16:29-16:30	2
peer_down	1
peer_up	1
Dec 19, 16:33-16:34	1
peer_up	1

NODE ID	
pca01ro1051bew	8
pca01ro1051zoi	8

PEER IP	
10.98.80.129	8
10.9[illegible]	8

REASON STR	
null	8
The remote system closed the session without a notifica...	8

BMP Peer State Change 64497:6



FORWARDING STATUS		
FORWARDED U...	46,449,485.38 Kbps	▼ 9%
UNKNOWN Uncl...	14,176,784.95 Kbps	▼ 5.6%
DROPPED Unknown	1,135.34 Kbps	▼ 2.5%
☑ DROPPED Adjacency	460.6 Kbps	▲ 270.4%
CONSUMED For us	218.95 Kbps	▼ 0.6%
DROPPED Policer	67.87 Kbps	▼ 72.6%
DROPPED bad TTL	19 Kbps	▲ 2614.9%
DROPPED RPF	1.65 Kbps	▲ new

PLATFORM ID		
ESC	203,349.62 bps	▲ 166.9%
HDB_SG	203,321.6 bps	▲ new
ECA	36,507.73 bps	▼ 11%
TV	27,124.21 bps	▲ 186.5%
EAC	573.72 bps	▲ 8.4%
MCC	551.4 bps	▲ new
DCSPLUS	116.2 bps	▲ 38.7%
BBLAB_GW	60.44 bps	▲ new

NODE ID		
sga01ro1011zoi	196.74 Kbps	▲ new
pcb05ro1010olt	59 Kbps	▲ 14710.6%
pca05ro1010zhh	58.63 Kbps	▲ 12976.9%
pcb05ro1020bew	49.73 Kbps	▲ 15.5%
pcb01ro1072olt	17.17 Kbps	▲ 5.2%
pca01ro1072zhh	16.65 Kbps	▲ 7.9%
pcb05ro1020olt	13.25 Kbps	▼ 8.4%
pca05ro1020zhh	13.11 Kbps	▲ 15.2%

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.

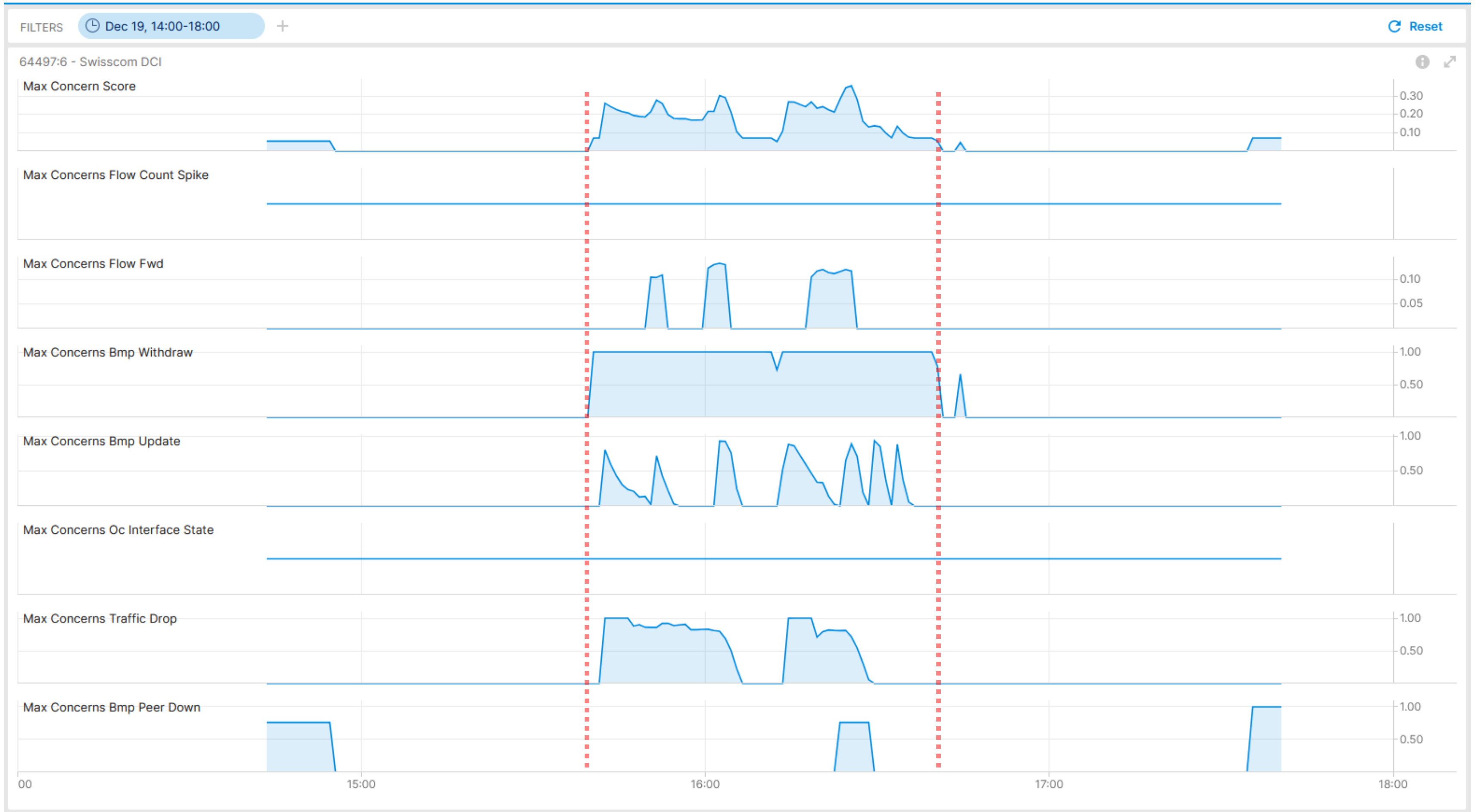
SOS

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

What Network Anomaly Detection observed, 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.**

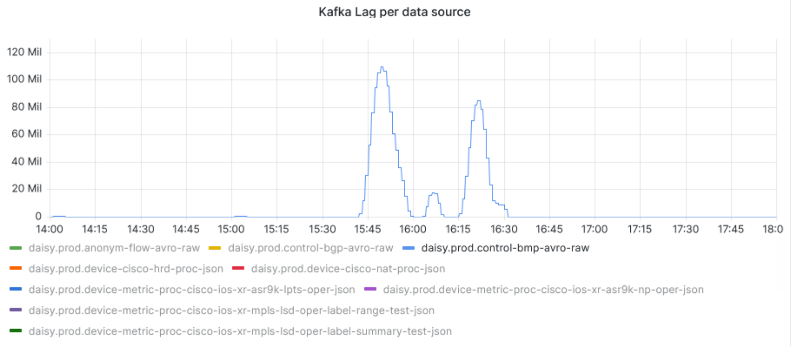
👍 **Missing Traffic recognized that connectivity is impaired.**

— 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 added in -02 revision the support of two reason code TLV's for prefixes crossing the warning and the maximum threshold.

draft-msri-grow-bmp-bgp-rib-stats authors added in revision -03 BMP statistics definitions describing how many routes until maximum prefix count has been reached.

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 data collection verified.

Maximum Prefix BGP Peer State Change

Want more?

- You are interested to see another Network Analytics Network Incident Postmortems? Please consider to attend SRv6OPS working group session on Tuesday 16:30 – 17:30.

- You want to contribute to the Network Anomaly Detection draft-ietf-nmop-network-anomaly-architecture and YANG to Message Broker Integration draft-ietf-nmop-yang-message-broker-integration and learn more? Please attend NMOP working group session on Tuesday 09:30 – 11:30, 18:00 – 19:00 for the hackathon related experiments or go onto the mailing list and contribute to the discussion.