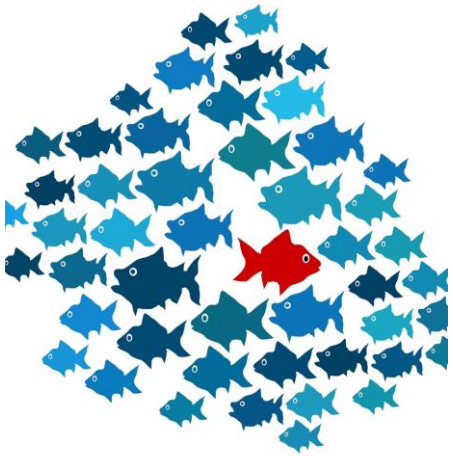
Network Telemetry and Big Data

From Flow Aggregation over Streaming Telemetry to Anomaly Detection

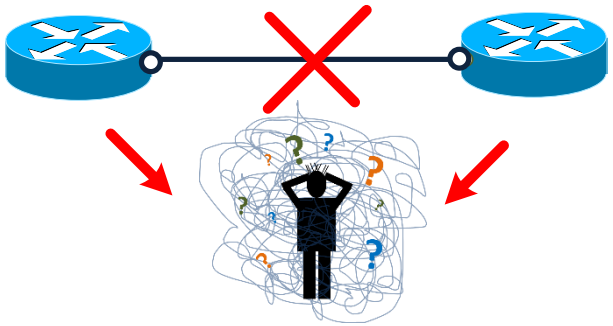
swisscom

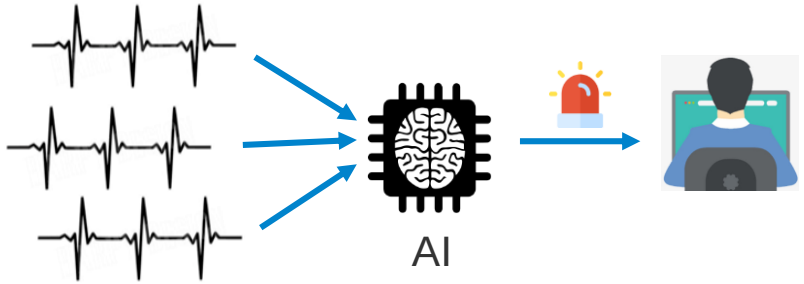
Towards Intelligent Monitoring

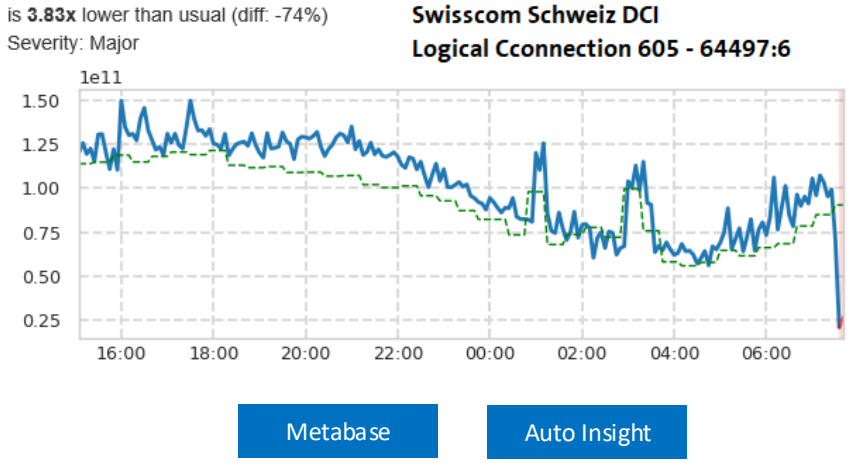
Devices do not know the network. Big Data & Analytics does



Bringing visibility and structure into what **appears to be chaos to humans**, but **makes sense so machines and networks**

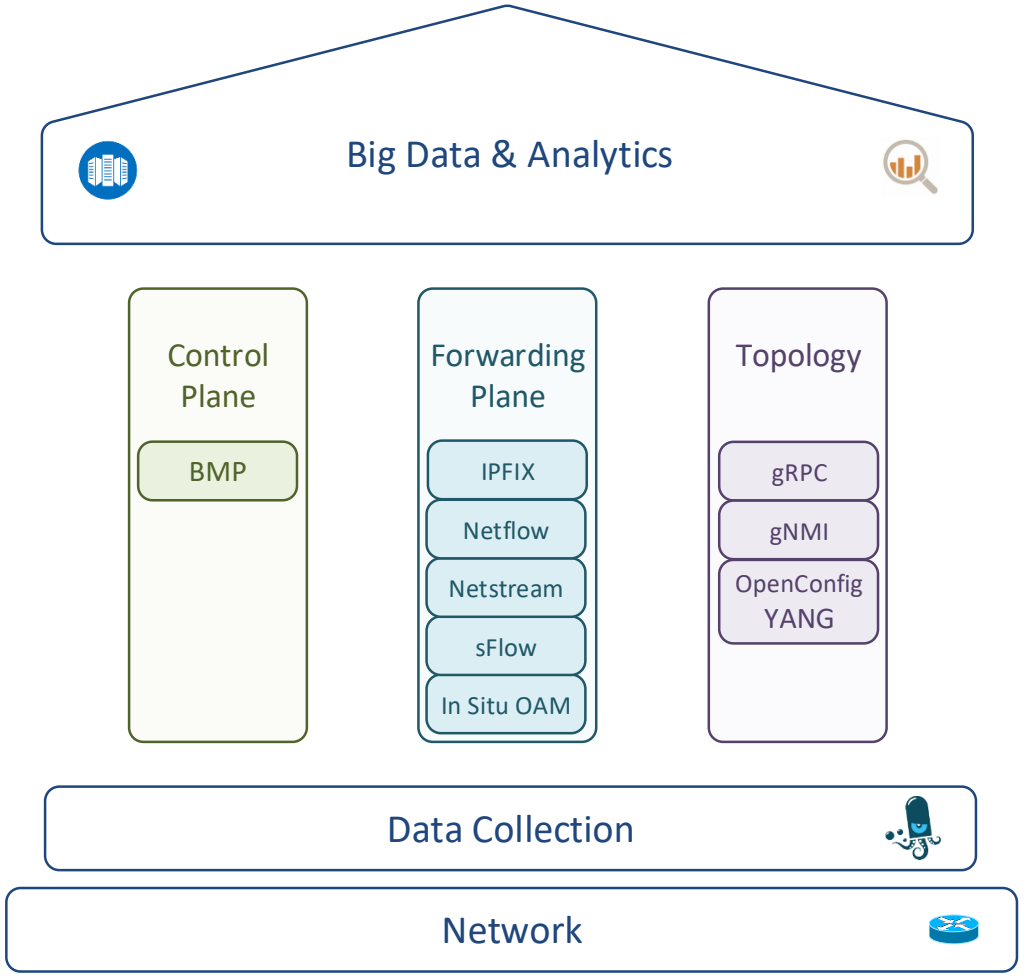






Data collection with Network Telemetry

Without order and structure, Big Data & Analytics gets lost



Network Telemetry

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

Topology metrics

- How logical and physical network devices are connected with each other.

Control-plane metrics

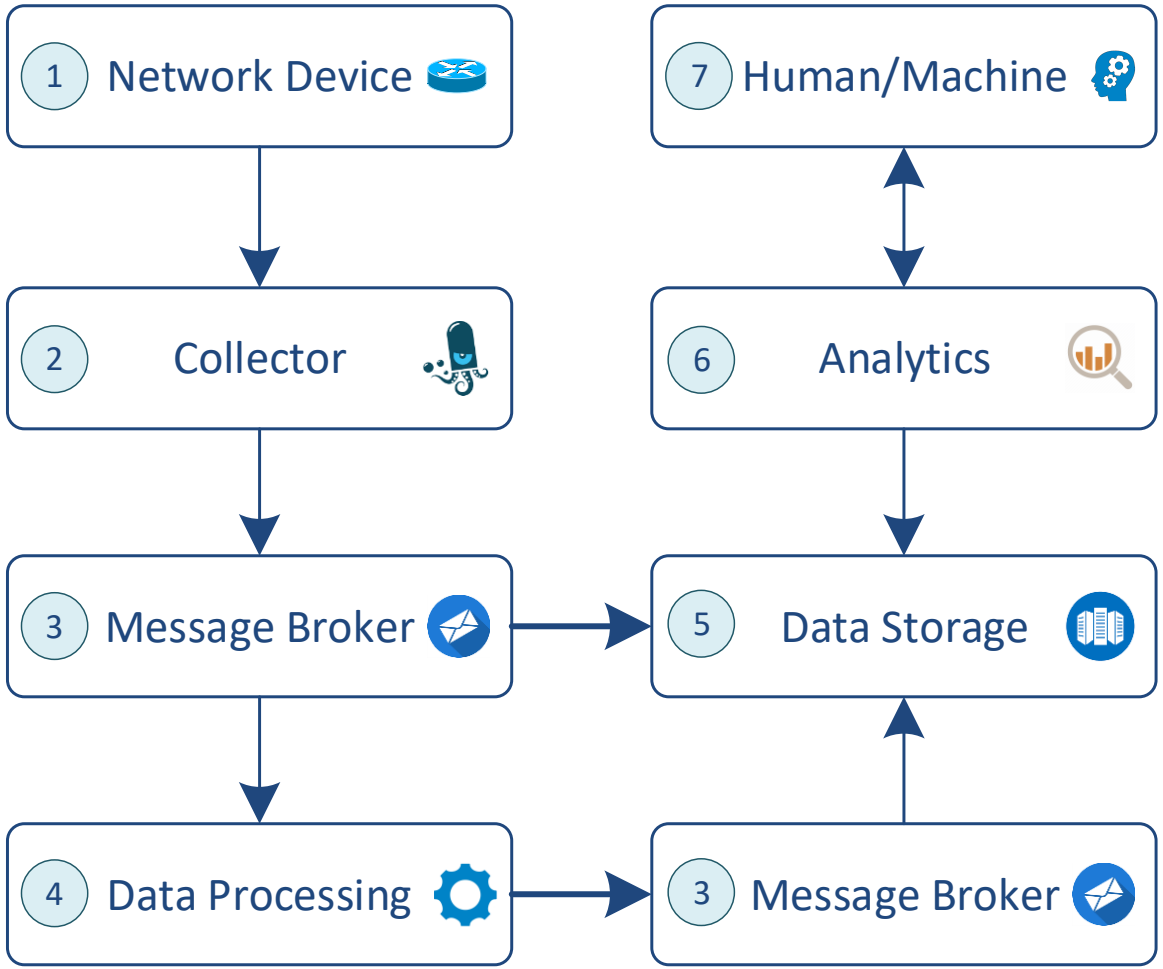
- How the network is provisioned and redundancy works.

Forwarding-plane metrics

- How traffic flows through the network.

Data pipeline with Big Data

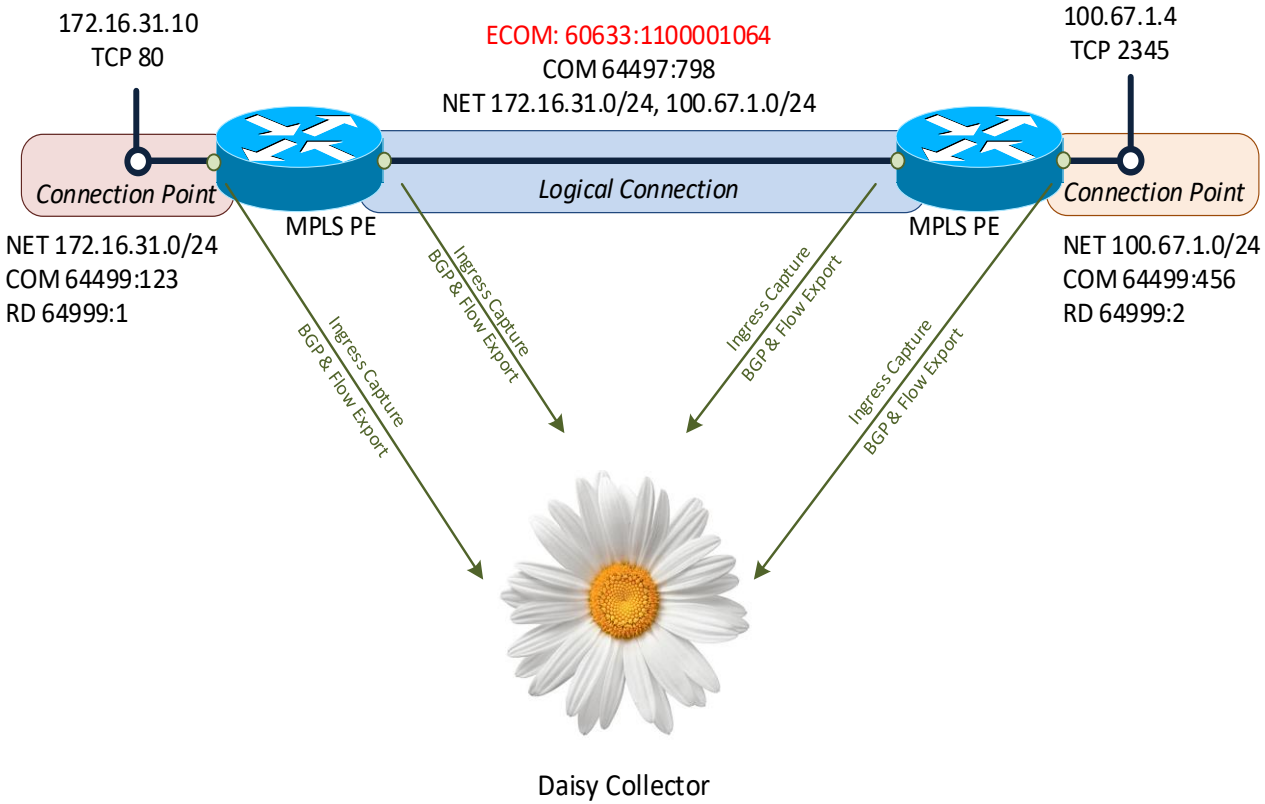
How do you eat an elephant? Piece by piece



1. **Pushing** metrics to collectors.
2. **Aggregate** or directly **ingest** into topics.
3. **Buffers, consolidate and forward.**
4. **Process** and re-ingest.
5. **Import** for midterm storage.
6. Access metrics and **perform baseline measurements.**
7. **Are informed** about events and possible service impact.

Intent Based Network - Insights into metric modelling

Without a BGP route-distinguisher, 192.168.1.1 isn't unique



Service Inventory consists of various logical elements which **influences predefined forwarding behavior within network**.

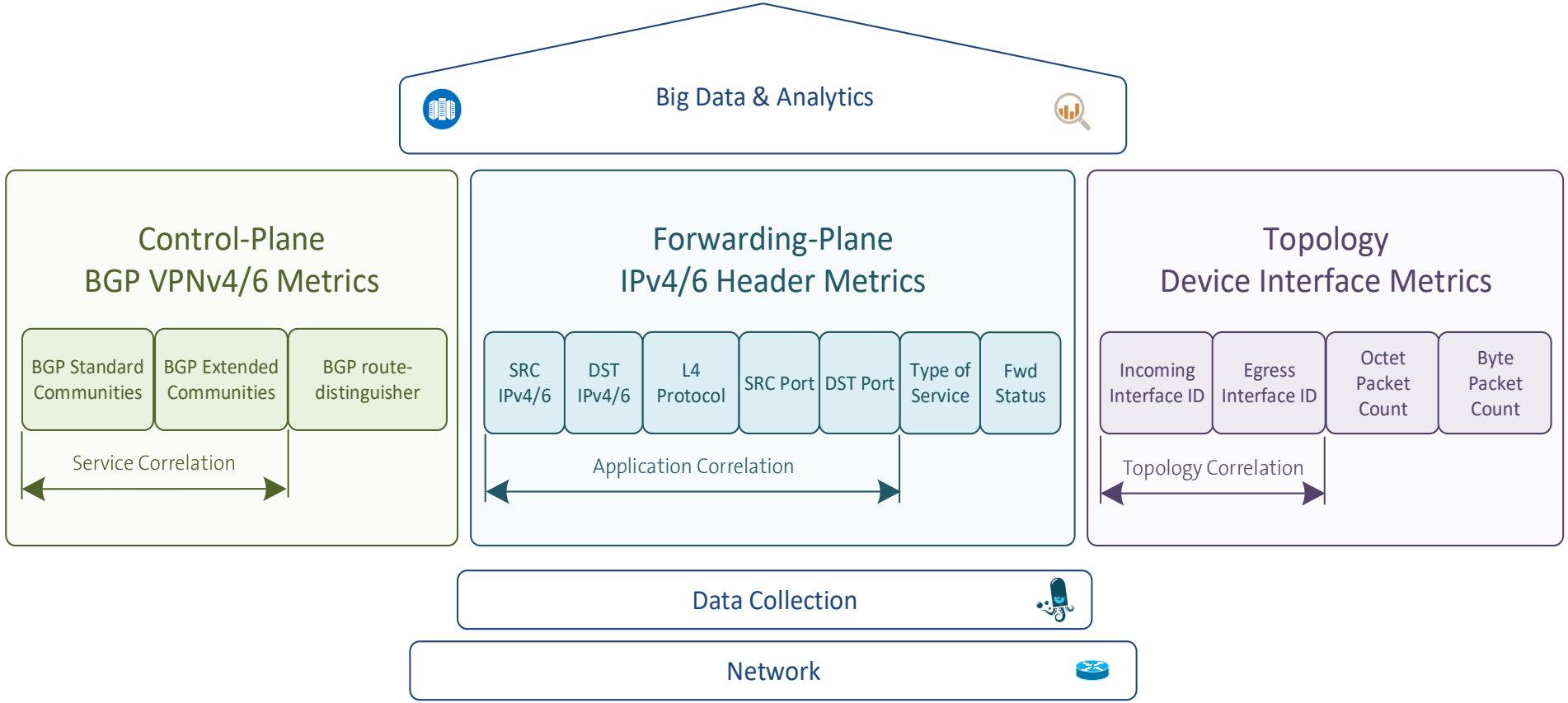
Service Inventory is **pushed with BGP** service and topology relevant **standard communities into BGP network**.

MPLS PE **exposing flow (forwarding-plane) and BGP (routing control-plane) to collector.**

Collector **collect, correlate, aggregate and ingest into message broker.**

VPN Service view with BGP VPNv4/6 Flow Aggregation

Depending our needs, we can look at data from different angles



BGP communities are used to **correlate to service inventory**

IPv4/6 address, Layer 4 port and protocol are used to **correlate to applications**

Interface ID will be used to **correlate to physical topology** collected through streaming telemetry

Kafka Message ingestion - BGP VPNv4/6 Flow Aggregation

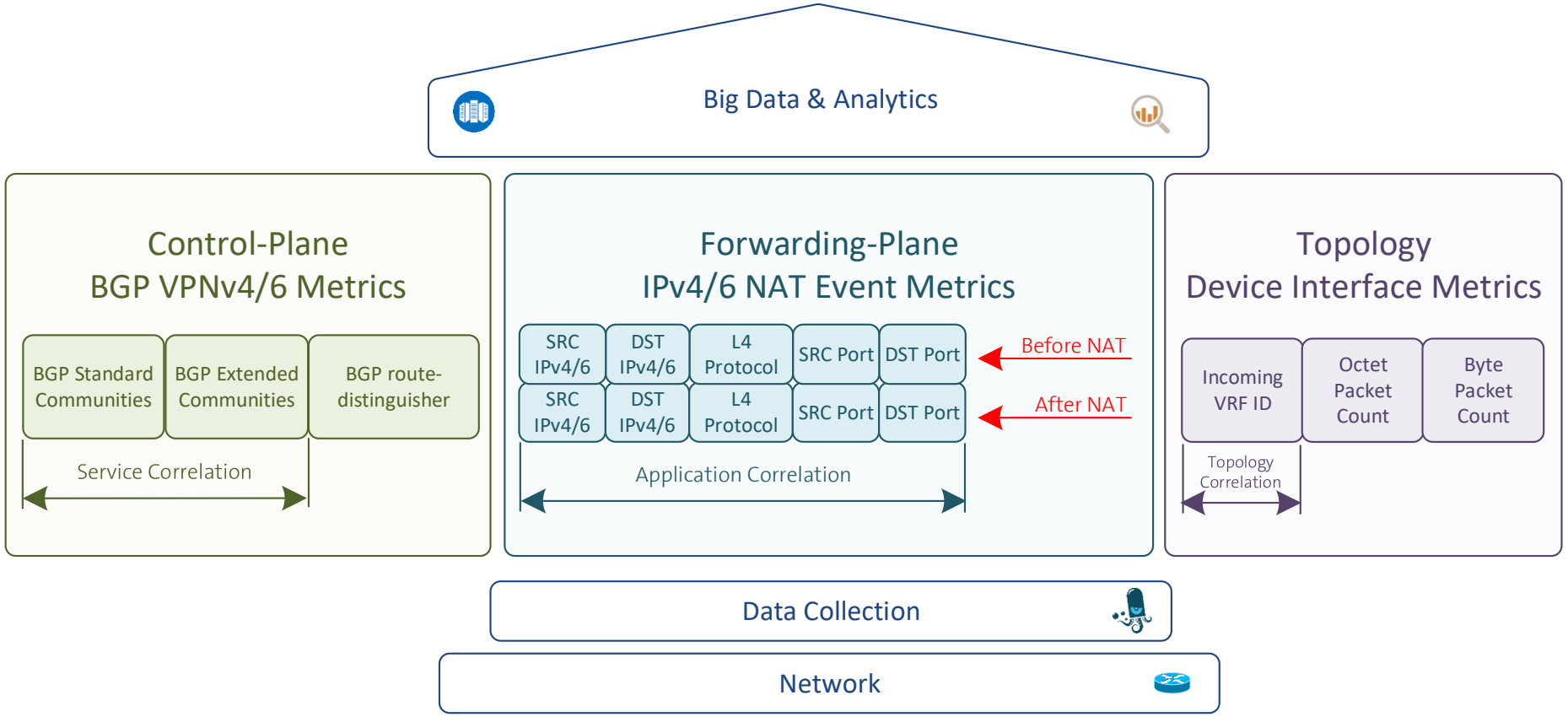
BGP and flow metrics come together into one message


```
{
    "event_type": "purge",
    "label": "sgs01ro1010olt",
    "comms": "60633:100_60633:265_60633:1001_60633:1032_64497:1528_64499:6000",
    "ecomms": "RT:12429:20000001_RT:60633:1100001715",
    "peer_ip_src": "138.187.57.53",
    "src_comms": "60633:100_60633:204_60633:1004_60633:1020_60633:1034_60633:10004_60633:10031_60633:10044",
    "src_ecomms": "RT:12429:30000001_RT:12429:32100001_RT:65511:1581_RT:65511:881581",
    "iface_in": 33,
    "iface_out": 47,
    "mpls_vpn_rd": "2:4200005685:11",
    "ip_src": "85.3.167.134",
    "net_src": "85.3.164.0",
    "ip_dst": "195.186.219.32",
    "net_dst": "195.186.219.32",
    "mask_src": 22,
    "mask_dst": 32,
    "port_src": 50379,
    "port_dst": 8443,
    "tcp_flags": "24",
    "ip_proto": "tcp",
    "tos": 0,
    "timestamp_start": "1540999270.0",
    "timestamp_end": "1540999270.0",
    "timestamp_arrival": "1540999295.307353",
    "timestamp_min": "1540999296.0",
    "timestamp_max": "1540999296.0",
    "in_iface_desc": "",
    "forwarding_status": "64",
    "vrf_id_ingress": "1610612736",
    "vrf_id_egress": "1610612752",
    "vrf_name": "",
    "stamp_inserted": "1540999260",
    "stamp_updated": "1540999299",
    "packets": 512,
    "bytes": 770048,
    "writer_id": "zhb01bgp01/10592"
}
```


- Example Kafka message containing correlated **BGP VPNv4, IPFIX Flow** and **collector** enriched metrics
- We use locally unique SNMP Interface ID (RFC 2863) for metric correlation on collector.

NAT service view with BGP VPNv4/6 Flow Aggregation

Giving insights in what is forwarded with which addresses at which point



BGP communities are used to **correlate to service inventory**

IPv4/6 address, Layer 4 port and protocol are used to **correlate to applications**

Interface ID will be used to **correlate to physical topology** collected through streaming telemetry

Kafka Message ingestion - BGP VPNv4/6 NAT Event Aggregation

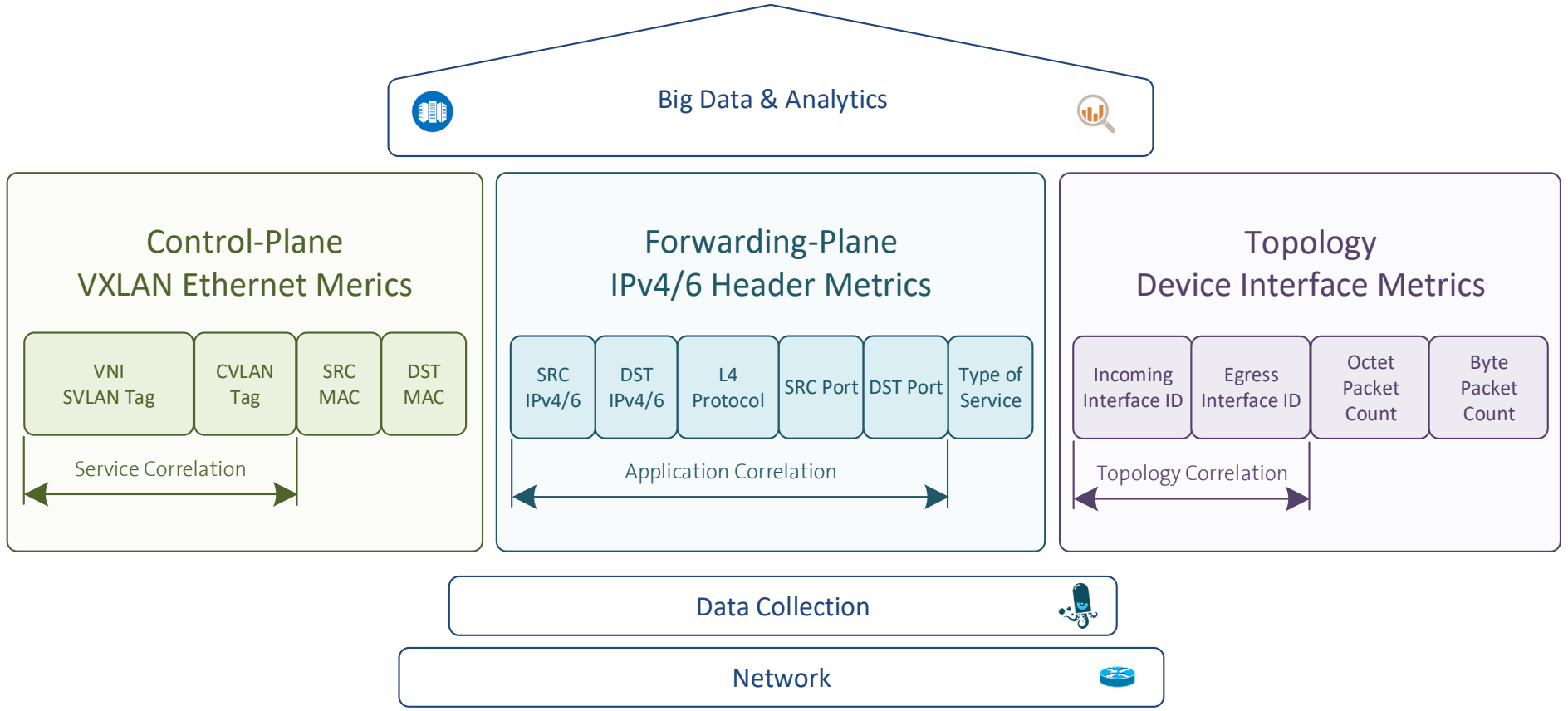
BGP and NAT event come together into one message


```
{
  "event_type": "purge",
  "label": "ipf-bew640-r-ss-01,SCB-NAT",
  "comms": "60633:299_60633:1001_60633:1033_60633:1111_64497:9998_64499:9013_64499:9014",
  "ecomms": "RT:60633:1100009988_RT:60633:1100009998_SoO:64499:1",
  "peer_ip_src": "138.187.57.59",
  "comms_src": "60633:299_60633:1001_60633:1034_64497:9998_64499:9016",
  "ecomms_src": "RT:60633:1100009998",
  "mpls_vpn_rd": "0:64499:1000990012",
  "ip_src": "10.100.100.2",
  "net_src": "10.100.100.0",
  "ip_dst": "8.8.8.2",
  "net_dst": "8.8.8.0",
  "mask_src": 24,
  "mask_dst": 24,
  "port_src": 17384,
  "port_dst": 17384,
  "ip_proto": "icmp",
  "post_nat_ip_src": "100.100.1.3",
  "post_nat_ip_dst": "8.8.8.2",
  "post_nat_port_src": 17384,
  "post_nat_port_dst": 17384,
  "nat_event": 1,
  "timestamp_start": "1556517591.666000",
  "timestamp_end": "0.000000",
  "timestamp_arrival": "1556517660.675699",
  "timestamp_min": "1556517682.000000",
  "timestamp_max": "1556517682.000000",
  "vrf_id_ingress": "1",
  "stamp_inserted": "1556517660",
  "stamp_updated": "1556517683",
  "writer_id": "daisy62bgp01/16124"
}
```


- Example Kafka message containing correlated **BGP VPNv4, IPFIF NAT Event** and **collector** enriched metrics
- We use locally unique ingress VRF ID for metric correlation on collector.

VPN Service view with VXLAN Flow Aggregation

Group by VNI or CVLAN to bring visibility into the forwarding path



SVLAN and CVLAN tag are used to **correlate to service inventory**

IPv4/6 address, Layer 4 port and protocol are used to **correlate to applications**

Interface ID is used to **correlate to physical topology** collected through streaming telemetry

Kafka Message ingestion – VXLAN Flow Aggregation

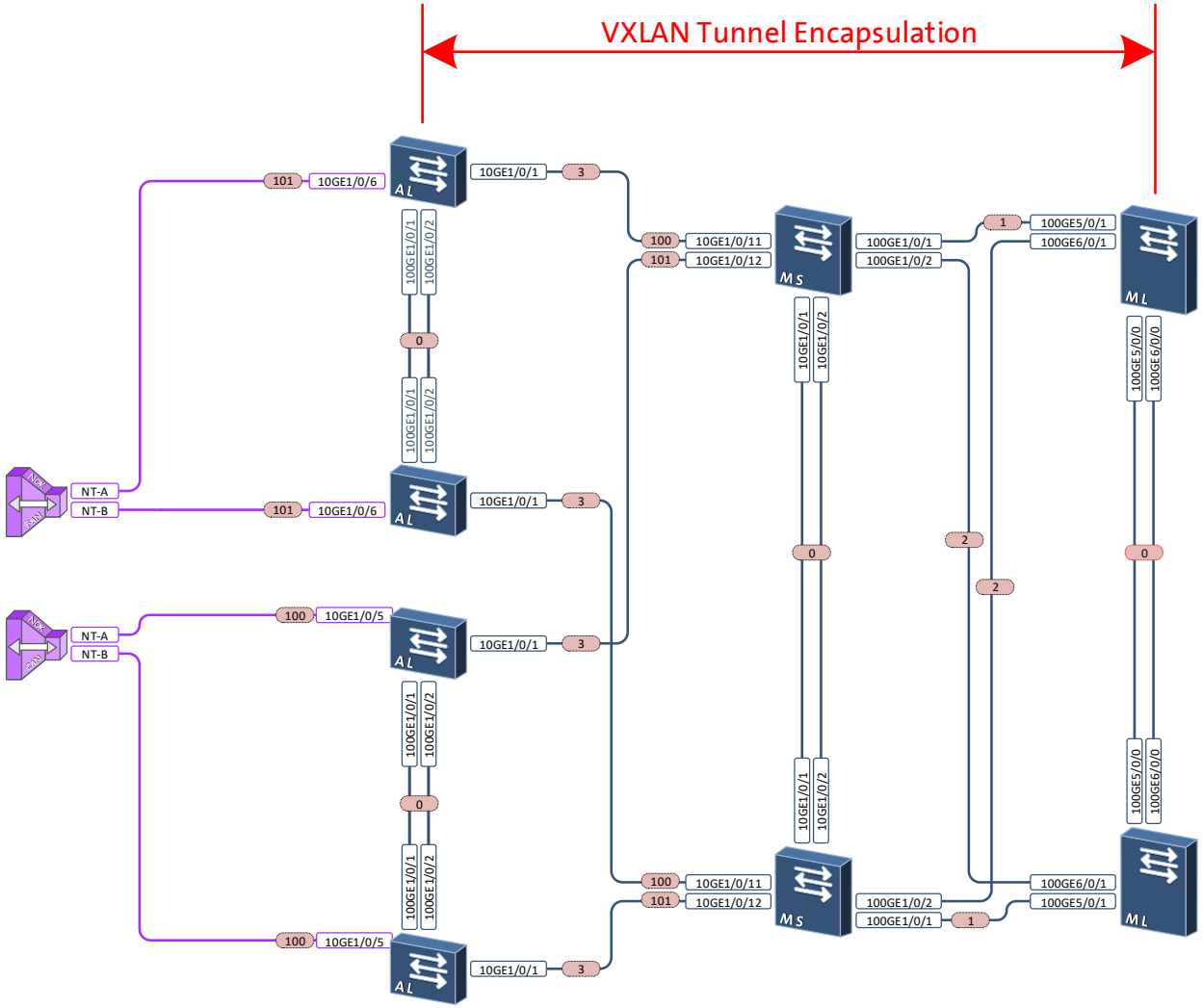
VXLAN and flow metrics come together into one message


```
{
    "event_type": "purge",
    "label": "ipi-zbb900-r-ms-01",
    "mac_src": "00:00:5e:00:01:01",
    "mac_dst": "00:70:01:00:00:2d",
    "peer_ip_src": "10.244.23.1",
    "iface_in": 0,
    "iface_out": 134,
    "ip_src": "10.239.16.1",
    "ip_dst": "10.238.19.144",
    "port_src": 67,
    "port_dst": 68,
    "tcp_flags": "0",
    "ip_proto": "udp",
    "tos": 112,
    "timestamp_start": "1541000396.0",
    "timestamp_end": "1541000396.0",
    "timestamp_arrival":
"1541000410.542835",
    "timestamp_min": "1541001726.0",
    "timestamp_max": "1541001726.0",
    "dot1qcvlanid": "104",
    "dot1qsvlanid": "100000000325d61",
    "stamp_inserted": "1541001720",
    "stamp_updated": "1541001727",
    "packets": 1,
    "bytes": 402,
    "writer_id": "daisyvxl02/67251"
}
```


- Example Kafka message containing correlated **VXLAN, IPFIX Flow** and **collector** enriched metrics
- Metric correlation is performed on routers.

Huawei and Swisscom Colaboration within Netcity

Alone we can do so little; together we can do so much

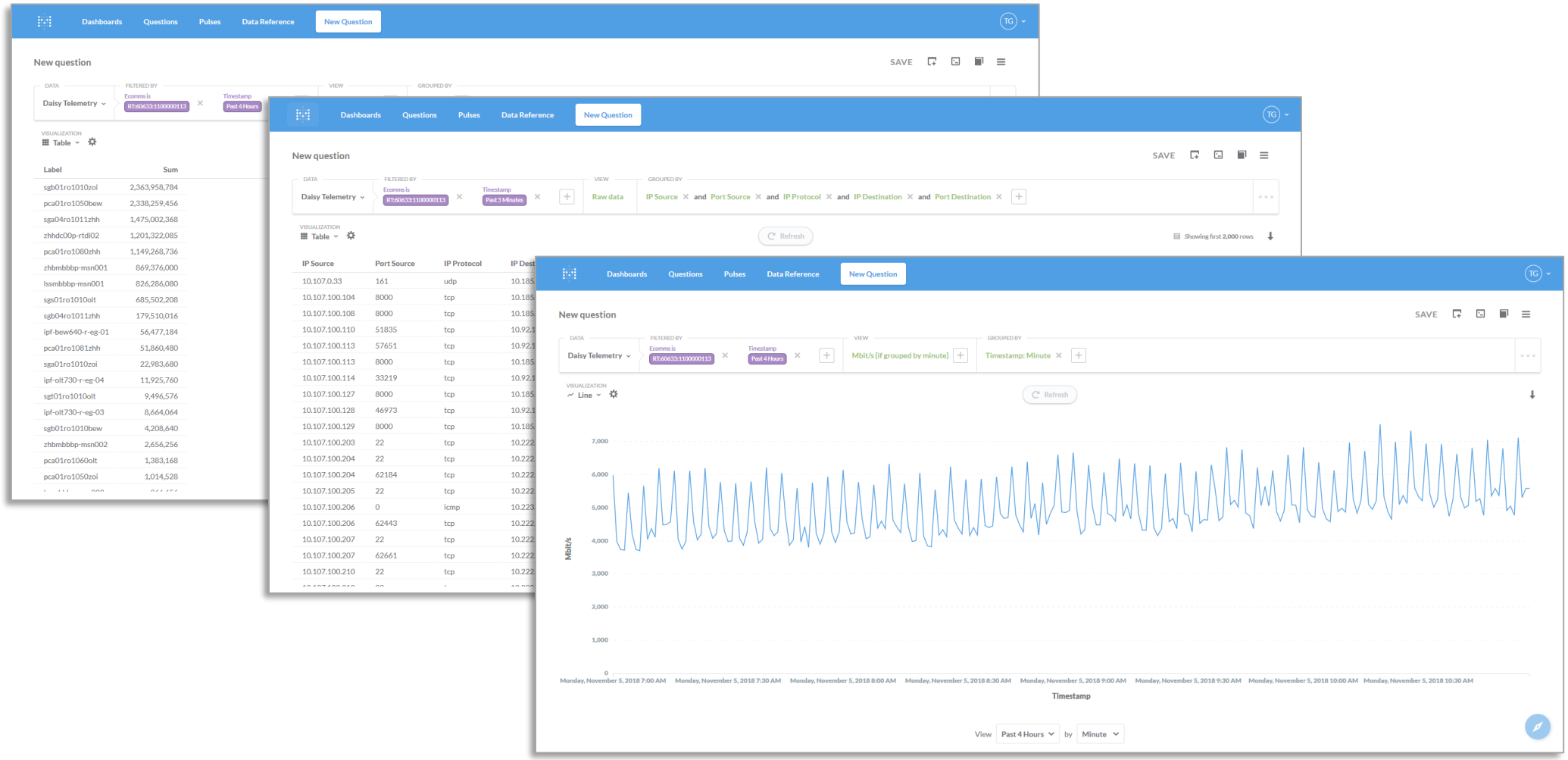


- **Since March 2018, Huawei and Swisscom collaborate under Netcity** to coordinate cutting edge technology developments.
- Within this scope, **Netstream VXLAN inner-IP feature was extended** to support the next generation Swisscom Broadband IPv6 underlay network and expose VXLAN header metrics on highly scalable Cloud Engine Platform in **latest VRP V200R005 release.**

HUAWEI

Its Demo Time !

Network Telemetry at Swisscom



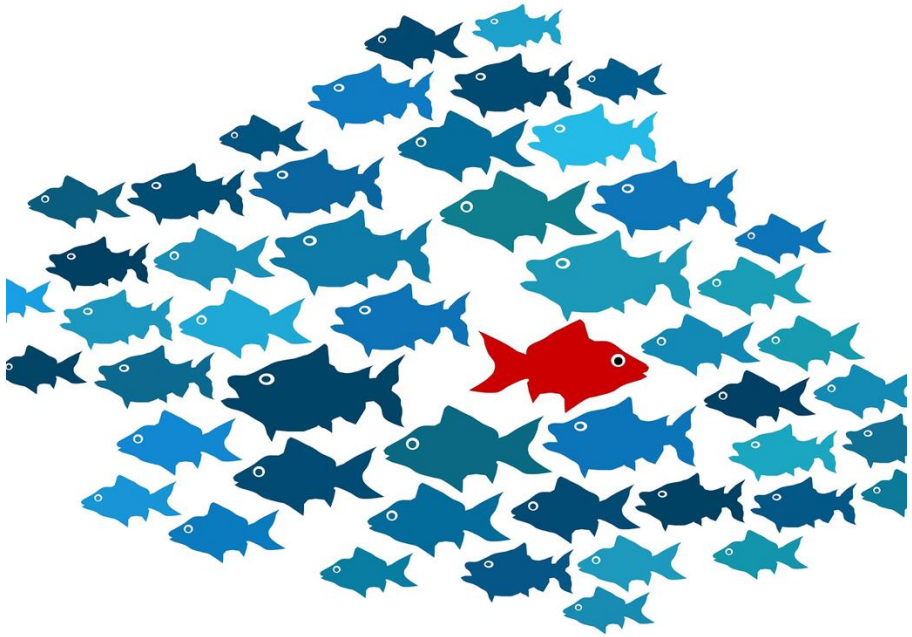
"Wait a minute, Doc. Are you telling me that you built a time machine... out of a DeLorean?!"

Marty McFly... 1985 - Back to the Future

Meerkat

Anomaly Detection Engine @ Timeseries Analytics Platform

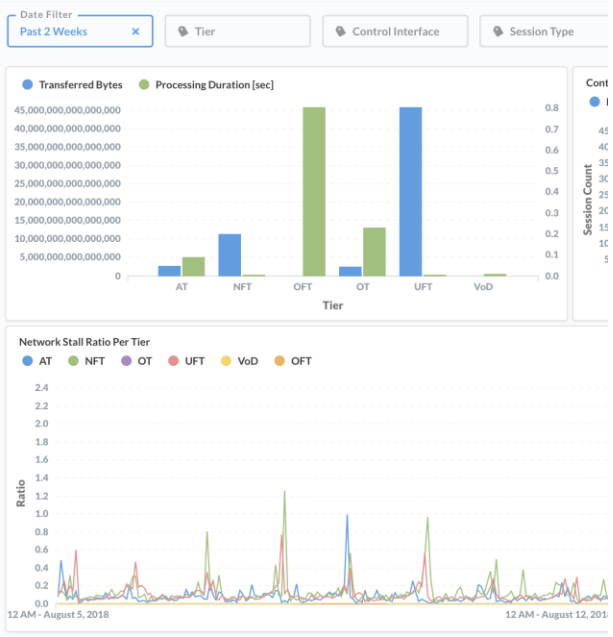
Zongren Liu
Senior Data Scientist

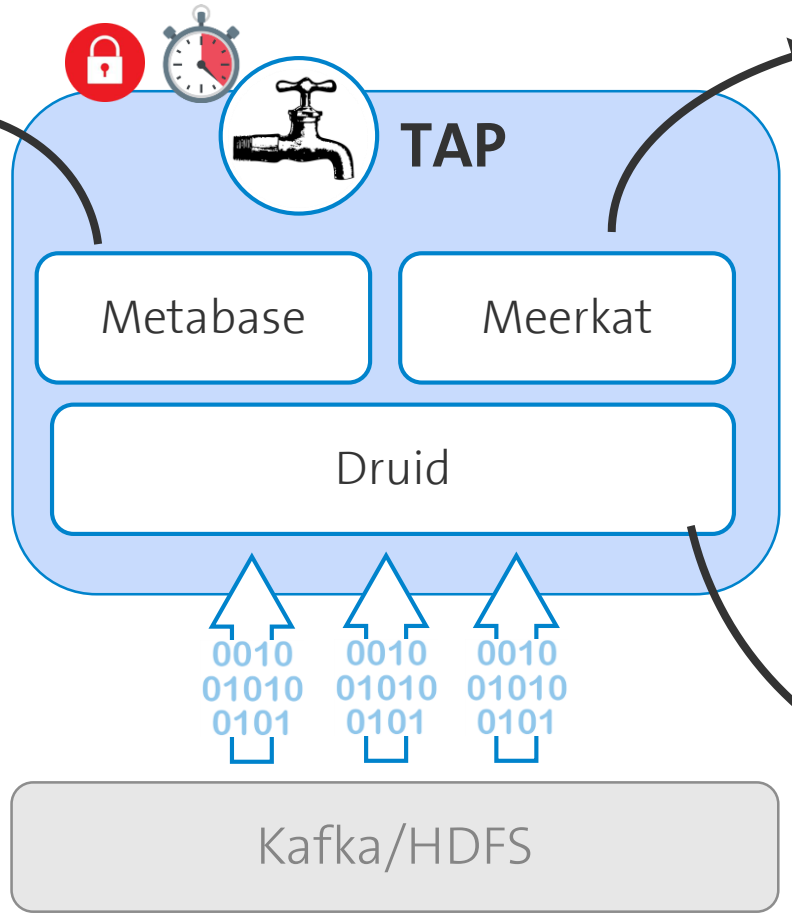


Timeseries Analytics Platform (TAP)

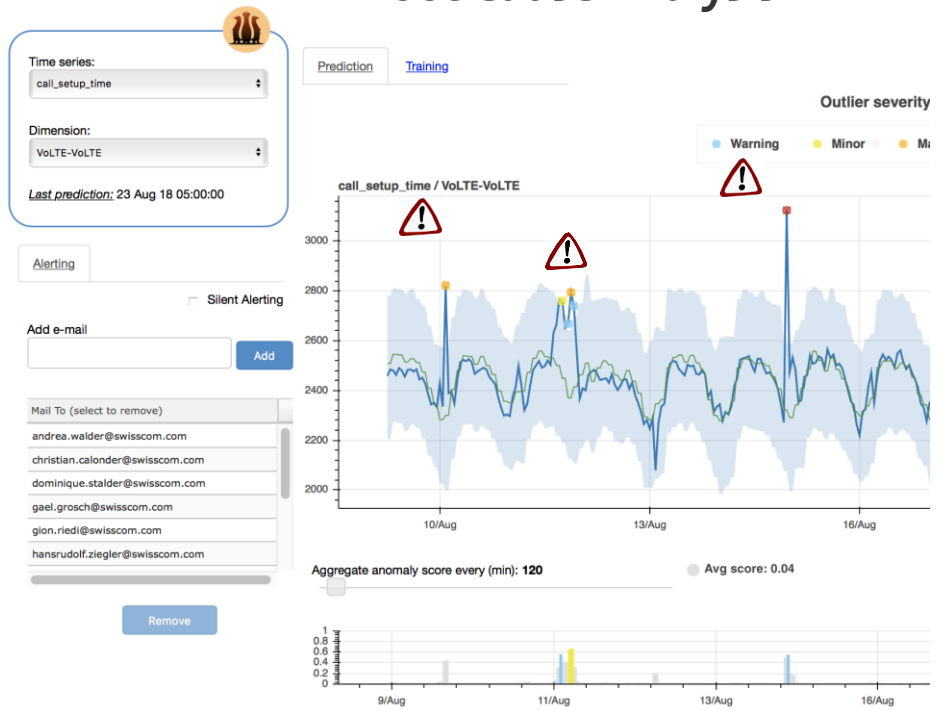
From Analytics to Root Cause Analysis

Self-service Analytics





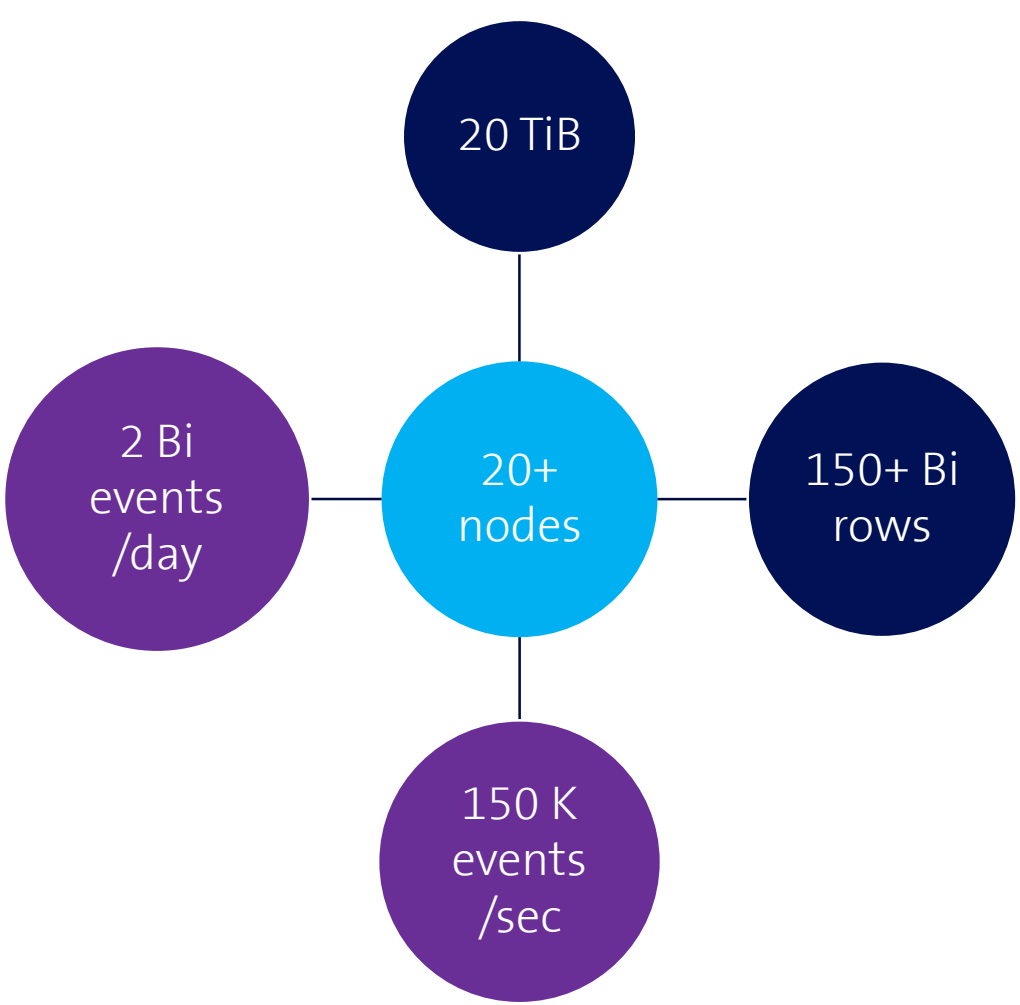
Anomaly Detection with Root Cause Analysis

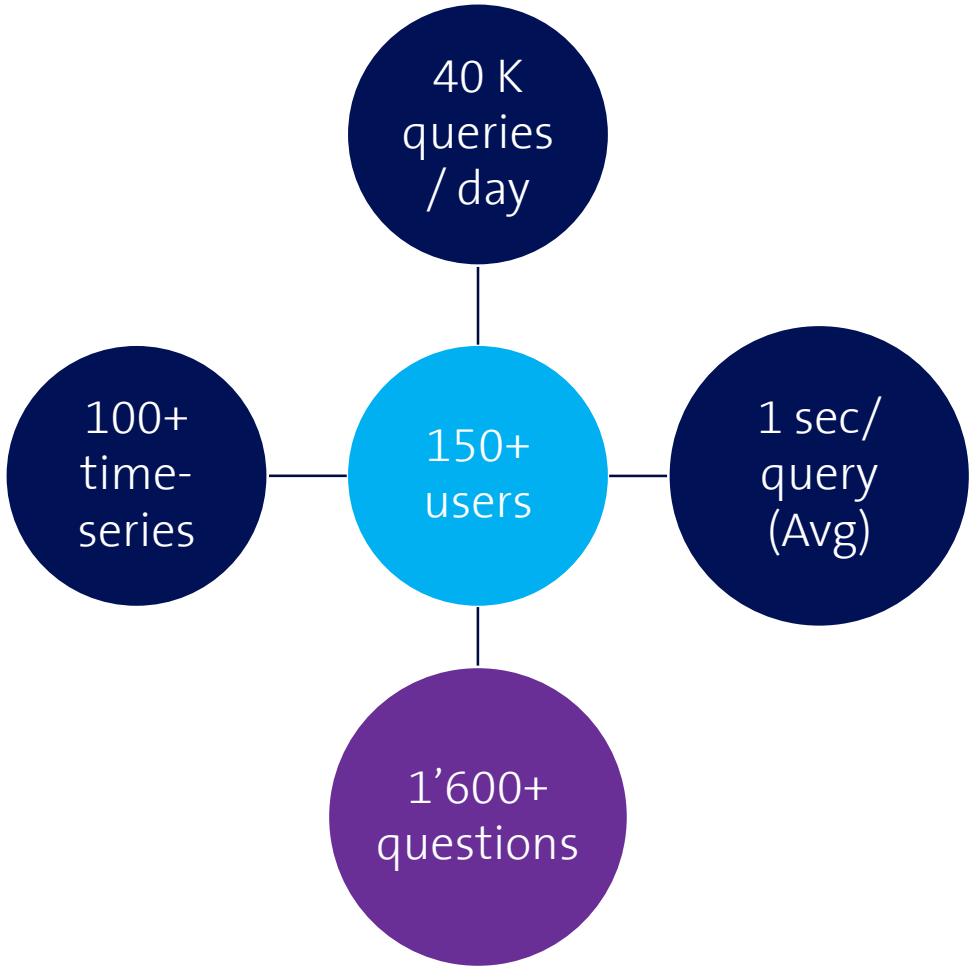


Realtime OLAP Storage & Query Engine

TAP in Numbers

In seconds to billions



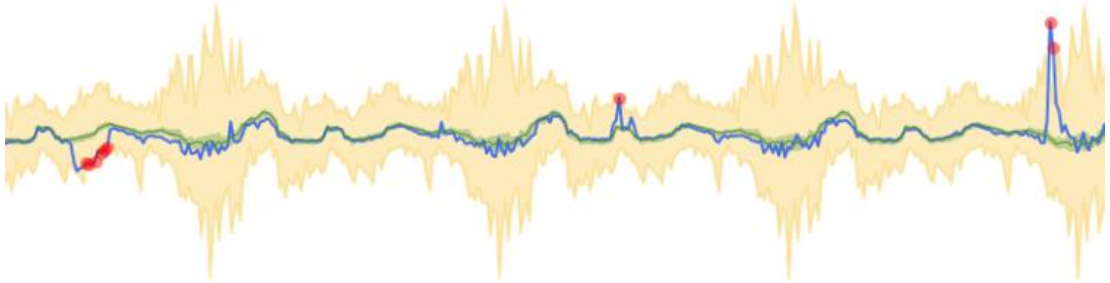


Towards Intelligent Monitoring

Threshold based approach is not enough

Need to build a statistical model that will

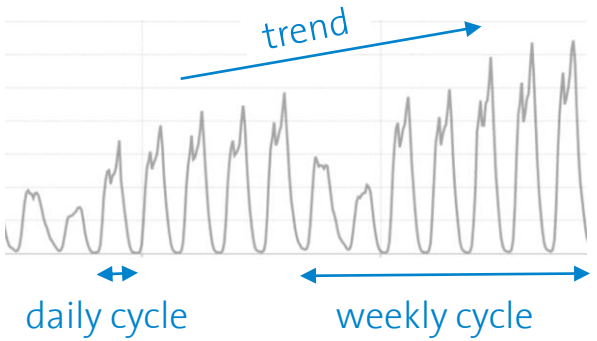
- Learn the time series pattern from historical data (and labels)
- Predict future values and compare with real measurements



- Outlier if : $|prediction - observation| > threshold$

Challenges

Need to be quick and accurate

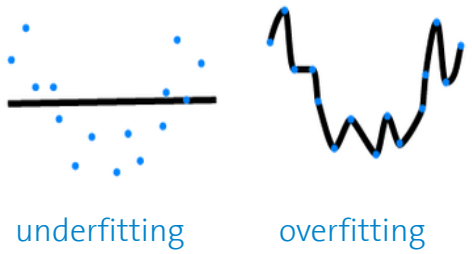


Realtime detection

- Online and fast algorithm
- Delay & data quality handling

No/few labels => ***Unsupervised (semi)*** learning

- Models auto selection/tuning



Root Cause Analysis

Understand the ***cause/impact*** of an alerted anomaly

Based on ***contextual data*** stored in Druid

Timestamp	Failed Calls	Cell Tower	Brand		Firmware	KPI 2
12:00	10	a1	b1	...	z1	70
12:00	300	a2	b1	...	z2	400
12:00	20	a2	b1	...	z3	8
12:00	70	a3	b2	...	z65535	23

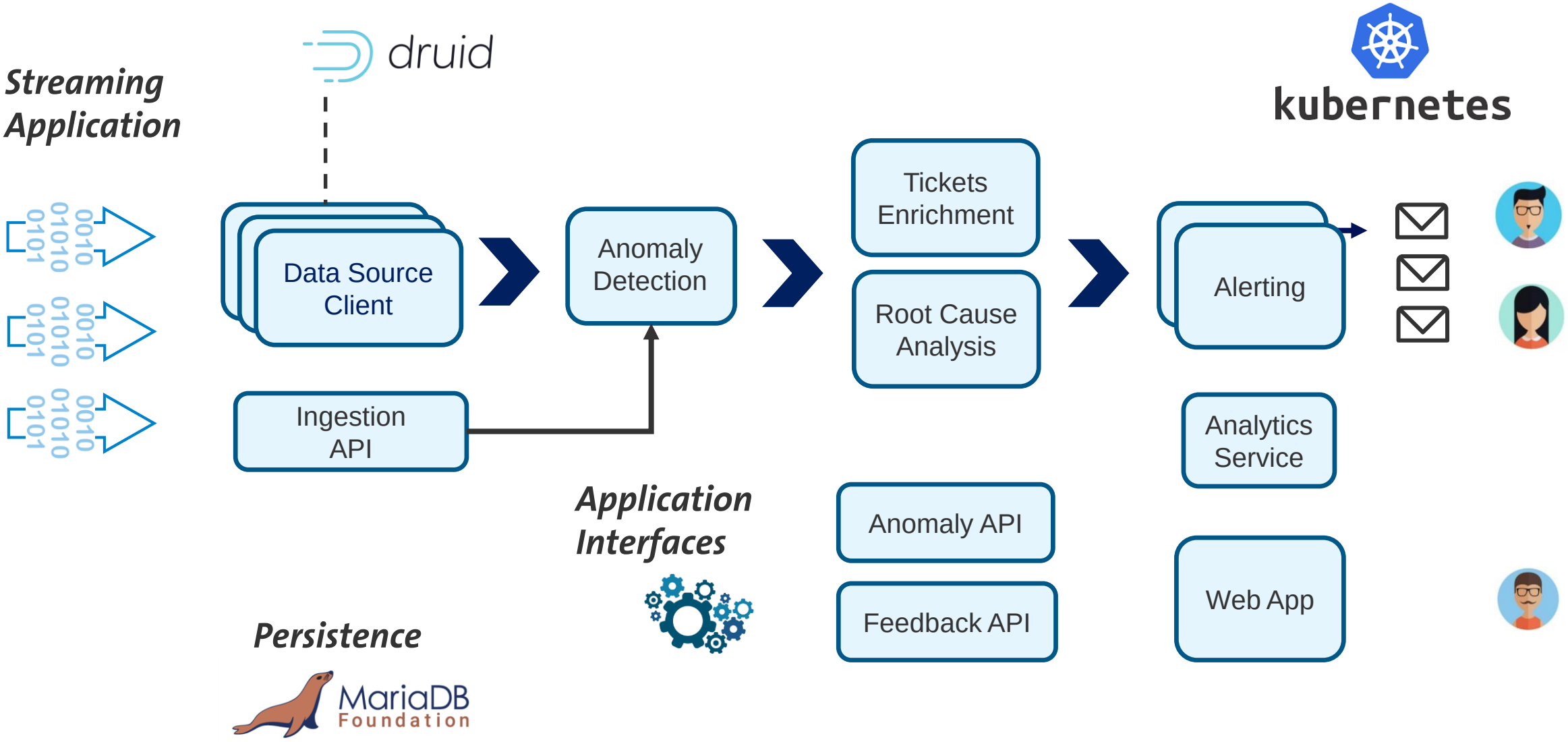
An engineering team may look at

- 6 different categories
- Having from 2 to 25’000 different values

Meerkat can do this a few seconds !

Meerkat – Building Blocks

How do you eat an elephant? Piece by piece



"Jesus Christ, Doc, you disintegrated Einstein!"

Marty McFly... 1985 - Back to the Future

whoami

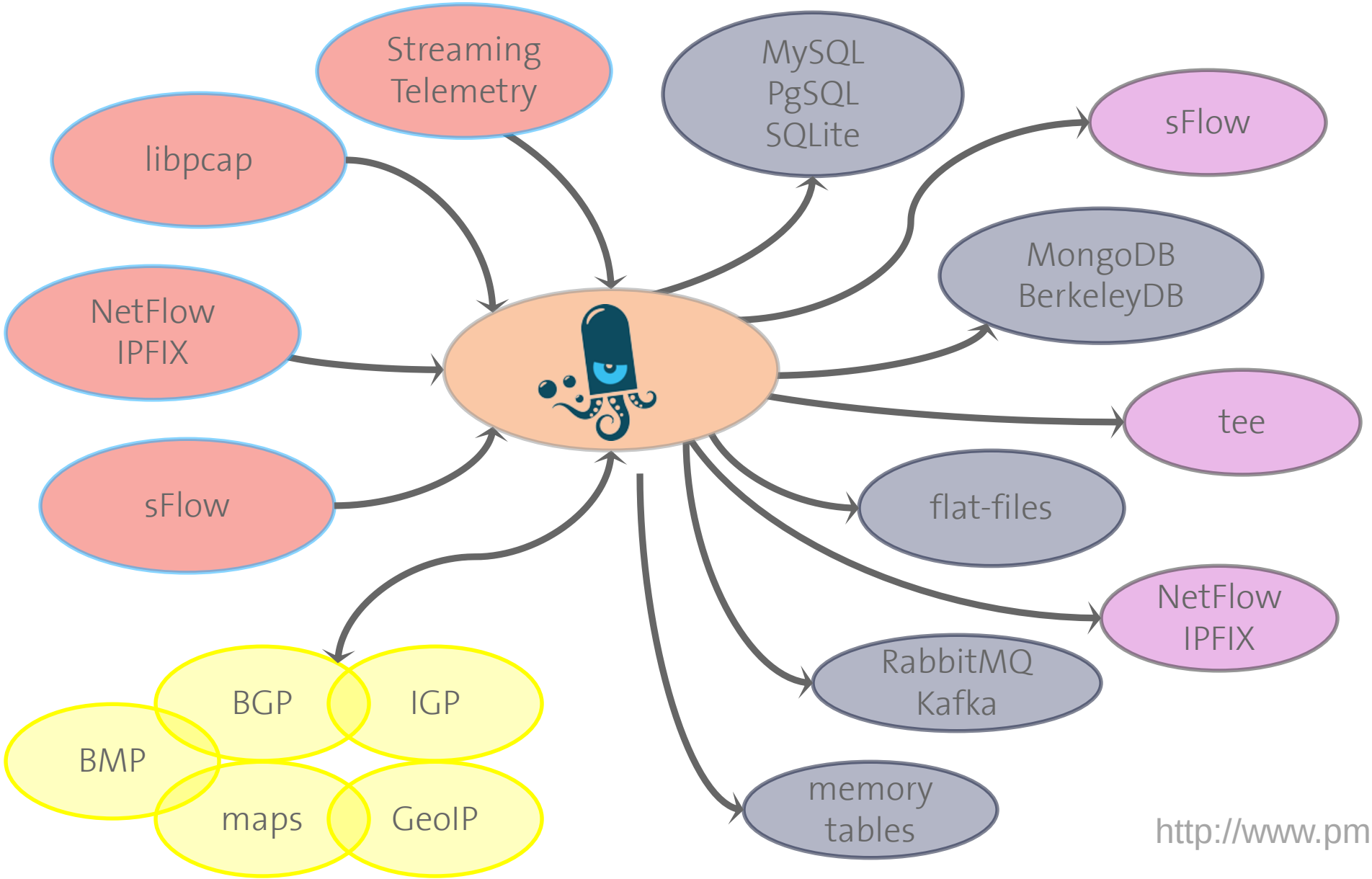
Paolo Lucente

GitHub: paololucente

LinkedIn: plucente

Digging data out of networks worldwide for fun and profit for more than 10 years

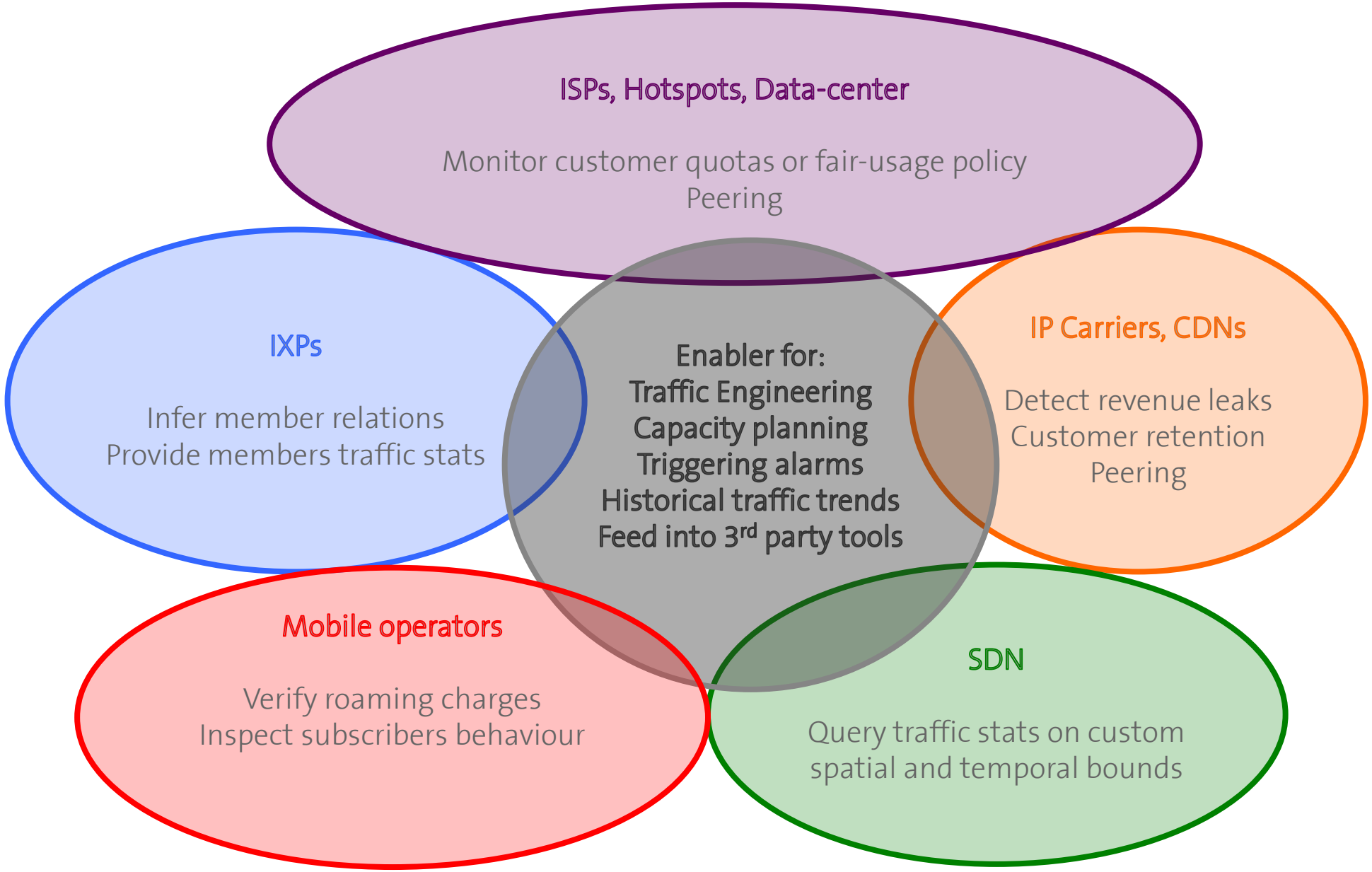
pmacct is open-source, free, GPL'ed software



http://www.pmacct.net

The use-case for message brokers

Use cases by industry



Some technical facts (1/2)

Pluggable architecture:

- Can easily add support for new data sources and backends

Correlation of data sources:

- Natively supported data sources (ie. BGP, BMP, IGP, Streaming Telemetry)
- External data sources via tags and labels

Pervasive data-reduction techniques, ie.:

- Data aggregation
- Filtering
- Sampling

Some technical facts (1/2)

Build multiple views out of the very same collected network traffic dataset, ie.:

- Unaggregated to flat-files for security and forensics; or to message brokers (RabbitMQ, Kafka) for Big Data
- Aggregated as [<ingress router>, <ingress interface>, <BGP next-hop>, <peer destination ASN>] and sent to a SQL DB to build an internal traffic matrix for capacity planning purposes

Enable analytics against the collected data sources (ie. BGP, BMP, Streaming Telemetry):

- Stream real-time
- Dump at regular time intervals (possible state compression)

Streaming Telemetry

A scalable replacement for SNMP:

- Push technology
- Subscribing to data of interest

A long journey to standardization ahead:

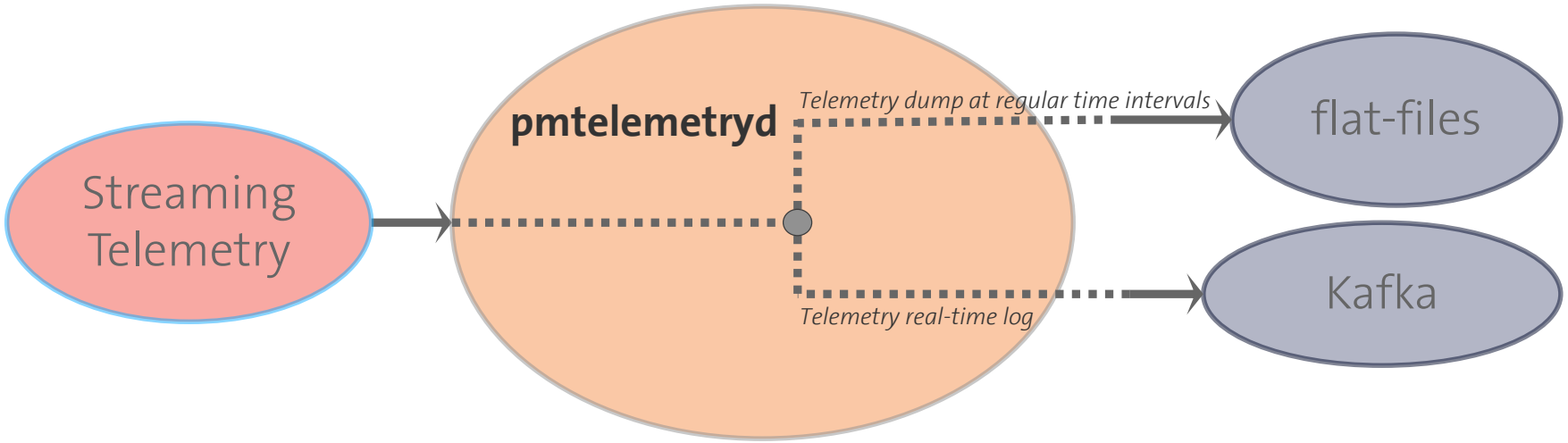
- Models: Openconfig and vendor-specific
- Transport: traditional, Netconf and gNMI
- RPC: Netconf (YANG Push) and gNMI
- Encoding: JSON and GPB

pmacct & Streaming Telemetry (1/3)

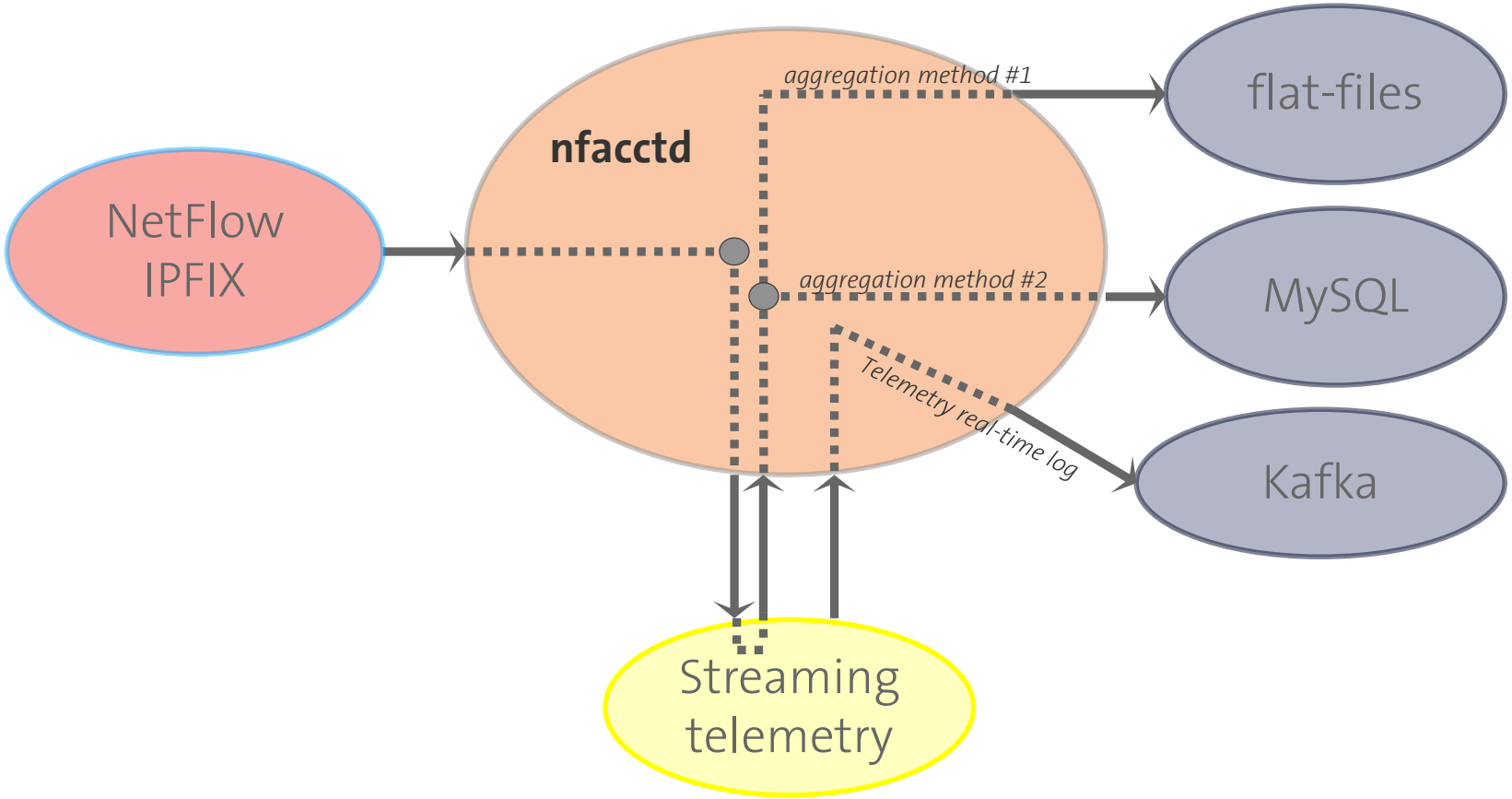
Mission statement:

- Integrate Streaming Telemetry data with other relevant technologies (ie. IPFIX, BMP, etc.)
- Especially in the current pre-standardization stage, offer an efficient multi-vendor collection layer for Streaming Telemetry

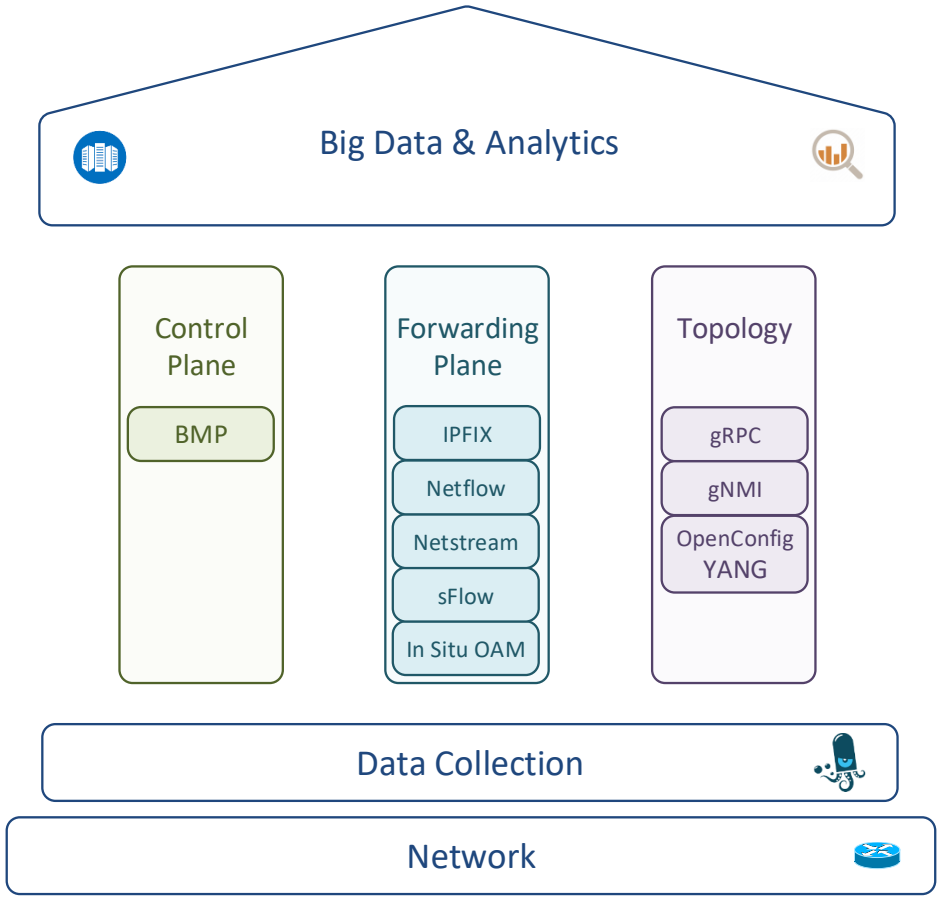
pmacct & Streaming Telemetry (2/3)

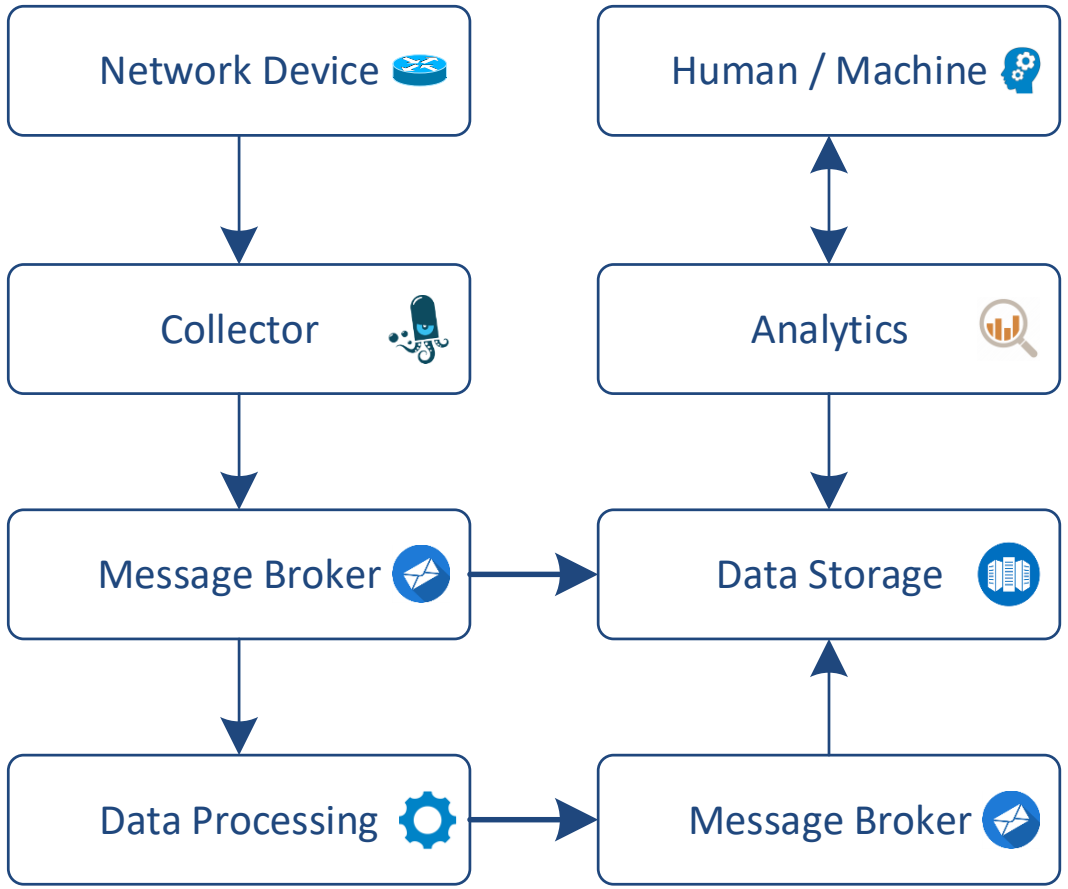


pmacct & Streaming Telemetry (3/3)



pmacct in Swisscom





Credits to: T. Graf (Swisscom) @ UBBF 2018

"I guess you guys aren't ready for that yet... But your kids are gonna love it."

Marty McFly... 1955 - Back to the Future

Credits

Paolo Lucente <paolo@pmacct.net>

Zongren Liu <zongren.liu@swisscom.com>

Thomas Graf <thomas.graf@swisscom.com>

Glossary
What?

pmacct	pmacct is a small set of multi-purpose passive network monitoring tools by Paolo Lucente
Daisy	Is the Network Telemetry platform name at Swisscom
Meerkat	Is the Anomaly Detection platform name at Swisscom
Netcity	Memorandum between Swisscom and Huawei to improve collaboration
Big Data	Is set of applications and databases to manage a very large amount of metrics
Anomaly Detection	Compares real time metrics to history by using artificial intelligence
Control Plane	Steers the network. Example: BGP, Border Gateway Protocol
Forwarding Plane	Forwards traffic through the network. Example: FIB, Forwarding Information Base
Topology	Physical connections between a group of network devices
Flow Aggregation	With topology or control-plane correlated Layer2-4 IP flow metrics
Streaming Telemetry	A SNMP replacement where topology metrics are streamed to a collector
OpenConfig	A consistent set of vendor-neutral data models written in YANG
Netconf	Network management protocol for device configuration
RabbitMQ, Kafka	RabbitMQ and Apache Kafka are the two most popular message brokers
Druid	Apache Druid is a high performance analytics data store for time series metrics

Glossary
What?

VXLAN	Virtual Extensible Local Area Network. An IP overlay encapsulation
VNI	Virtual Extensible LAN Network Identifier. A unique VPN tunnel identifier
IPFIX	IP Flow Information Export. A protocol to collect IP flow metrics
BMP	BGP Monitoring Protocol. A protocol to collect BGP control-plane metrics
BGP	Border Gateway Protocol. The routing protocol used in the Internet and at large networks
ASN	Autonomous System Number. A network domain used in the routing protocol BGP
IGP	Internal Gateway Protocol. Providing the next-hop attribute for BGP.
gRPC	Google Remote Procedure Call, transport Protocol for Streaming Telemetry
gNMI	Google Network Management Interface, IETF Draft for Streaming Telemetry
YANG	Yet Another Next Generation. A data modelling language for topology metrics
JSON	JavaScript Object Notation. A lightweight data-interchange format
GPB	Google's language/platform-neutral, extensible mechanism for serializing structured data