Broadband network telemetry

Efficient data telemetry sets the foundation for automation in broadband access

White paper

With the arrival of SDN in the broadband access network, the opportunity arises for advanced automation to improve network operations and unlock new functionality that decreases costs and creates revenue opportunities. Key to this opportunity are the vast amounts of network data to be collected, processed and stored. A structured approach to the telemetry that feeds this data for analysis provides the foundation for successful automation.

Contents

Introduction

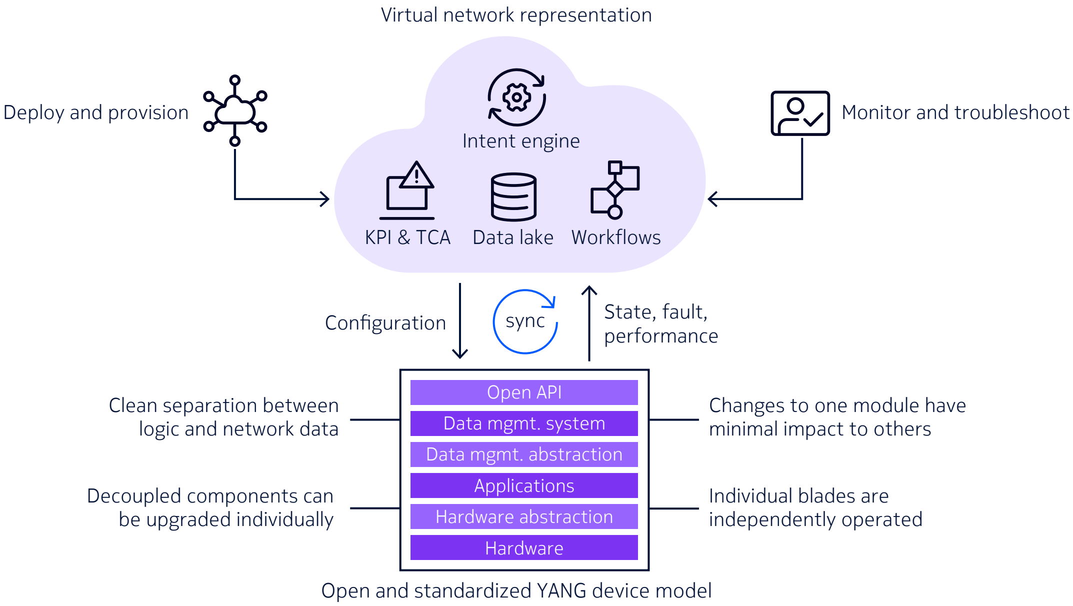
Automation is a growing business imperative for increasing efficiency when operating modern broadband networks. Pivotal in unlocking these capabilities is the availability of high-quality telemetry data feeding applications such as analytics, automated troubleshooting, and AI/ML. This means seeing telemetry as much more than bulk data collection at regular intervals (typically 15 minutes), and conventional Operations, Administration and Maintenance (OAM) techniques and protocols. Moreover, traditional pull-based data collection does not scale to the needs of increased automation.

Modern push-based streaming telemetry systems have enhanced capabilities for better flexibility, scalability, accuracy, coverage and performance for remote collection, correlation, and consumption of data, allowing for automated control using data-driven decision making and closed-loop automation. The access to real-time data which is structured and ready for use by systems performing data analytics is an enabler for use cases such as network automation, traffic optimization, anomaly detection, automated troubleshooting, and preventive care.

Network evolution to SDN and cloud-native designs

Software-defined networking (SDN) in access networks creates an open and modular disaggregated cloud-native architecture. As networks are controlled by software functionality, operations can be more easily automated using open standardized APIs.

Figure 1 – Software-defined access network architecture


Virtual network representation
Deploy and provision
Intent engine
Monitor and troubleshoot
KPI & TCA
Data lake
Workflows
Configuration
sync
State, fault, performance
Clean separation between logic and network data
Decoupled components can be upgraded individually
Open API
Data mgmt. system
Data mgmt. abstraction
Applications
Hardware abstraction
Hardware
Changes to one module have minimal impact to others
Individual blades are independently operated
Open and standardized YANG device model

Through considered rearchitecting, software and hardware of the traditional node architecture are disaggregated. The transparent APIs, the modular node software, the decoupling of ONT management and the flexible YANG-based device modeling transform the access network into a programmable and flexible infrastructure. Individual blades are independently operated, changes to one software module have minimal impact to others, and decoupled components can be upgraded individually.

Standardized interfaces for management, control and orchestration are vital to effectively automate. The SDN management and control interfaces of access and edge functionality are defined by Broadband Forum's CloudCO. It defines an extensive set of YANG data models for configuration, operational state and event notifications. By implementing CloudCO models in SDN controllers, operators get a single-pane-of-glass view to visualize, optimize, and enhance the fixed access network through open APIs across a multi-vendor, multi-technology environment.

In this data-centric cloud architecture, the SDN controller creates a virtual network in the software layer: a digital representation of the physical network, containing configuration, state, and performance data. The authoritative source of data is in the cloud and no longer in the individual network elements. All configuration, performance and diagnostic data, including logs and alarms, is centralized in a common data lake. For example, changes to the node configuration are applied directly in the cloud and the nodes take the cloud configuration via an online synchronization process. The SDN controller does not only store data, but also offers streaming on the north-bound APIs to allow other applications to pull data from the data lakes. The access to big volumes of centralized data is vital for analytics and sets the foundation for machine learning, which broadens greatly the set of network behavior that can be captured:

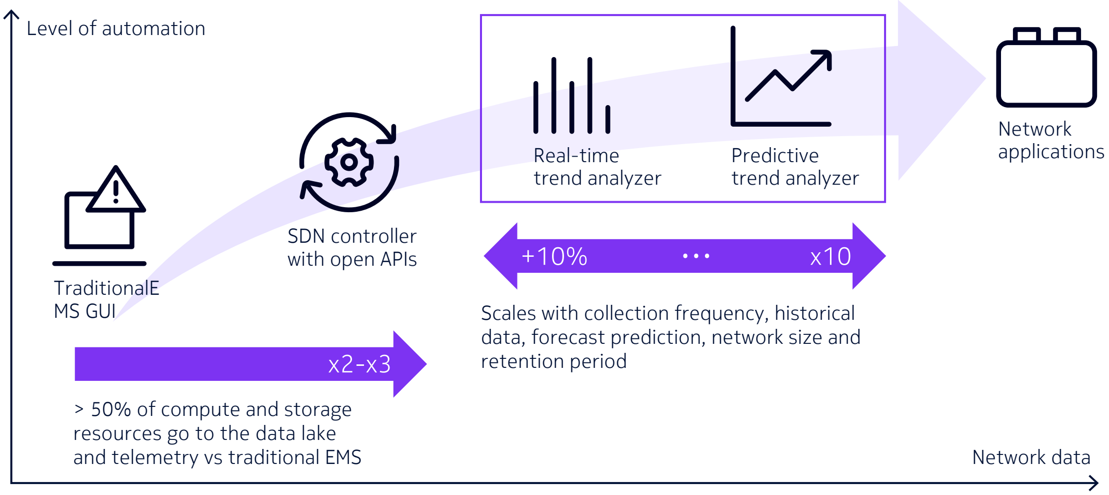
- **Size**, behavior that requires processing large amounts of data.
- **Speed**, behavior that requires real-time actions.
- **Complexity**, models that provide insights into correlation and causation.

Scaling up the level of automation incrementally

In cloud-native network architectures, the focus for data insights shifts from the node to the controller.

In the past, the node only reported data every 5 or 15 minutes. It was also the node that did most of the heavy lifting in data aggregation and metric computations. When large volumes of data were requested, device performance bottlenecks led to data breakpoints. As a result, network monitoring was limited in its use and scalability: it restricted the number of metrics which could be collected and did not support real-time monitoring. The polling caused higher CPU load and might even affect the proper functioning of the access equipment when polling all possible parameters. The traditional network management system (NMS) also had limited intelligence: it collected the data stored in the node, visualized it in its GUI, and didn't process incoming data or perform historical storage of collected data.

Figure 2 – The evolution from traditional NMS to SDN controllers with smart network applications


Level of automation
TraditionalE
MS GUI
SDN controller
with open APIs
Real-time
trend analyzer
Predictive
trend analyzer
Network
applications
+10%
•••
x10
Scales with collection frequency, historical data, forecast prediction, network size and retention period
x2-x3
> 50% of compute and storage resources go to the data lake and telemetry vs traditional EMS
Network data

An SDN controller differs from a traditional NMS. Thanks to push-based streaming telemetry, SDN solutions are capable of monitoring more network devices and process more parameters: 10x-20x more counters compared to legacy NMS solutions. By sending and processing data as it becomes available, it is also able to provide a real-time view of the network. It is also up to the SDN controller to optimize the collection model and define the rules to act upon the incoming data. The SDN controller has a range of flexible options to define KPIs, set threshold crossing alarms, trigger network policies, start automated workflows, or perform closed-loop optimization.

As the level of automation increases, the footprint of the solution will also increase: more than 50% of the compute and storage resources go to scaling the data lake and embedded telemetry systems. As more use cases and analytics applications are added, it is very important to ensure telemetry systems are properly designed in line with business needs. The volume of network data that can be obtained can be huge and the system load increases quickly with collection frequency, historical data, forecast prediction, network size, and retention period.

Data governance is the foundation for automation

To achieve a high level of automation, data should be structured and ready for use by systems performing data analytics. This means data governance becomes the most valuable component of the whole network automation system. The capture, transmission, storage and processing of massive data sets should be translated into hardware requirements and embedded in software algorithms.

It is good to realize how much data is commonly needed to implement modern data telemetry systems. In many cases, we are talking about millions of data points. At this scale, duplication of data acquisition, pre-processing and storage is to be avoided. When data is not easily consumable via open APIs, multiple teams and applications may duplicate data in isolation with limited consistency or re-use. It also leads to retaining aged information, leading to exploding compute and storage costs.

There are several standards that are dealing with network telemetry such as IETF RFC 9232 - Network Telemetry Framework (NTF), ETSI Zero Touch Network and Services Management (ZSM), Generic Autonomic Networking Architecture (GANA), BroadBand Forum TR-436 and WT-486 and TMForum. Among them IETF NTF is focusing on telemetry systems themselves. Others, such as BBF TR-436 and ETSI ZSM, are looking at telemetry in the context of Automated Intelligent Management and Analytics & Machine Learning scenarios).

Modern network data telemetry systems should be:

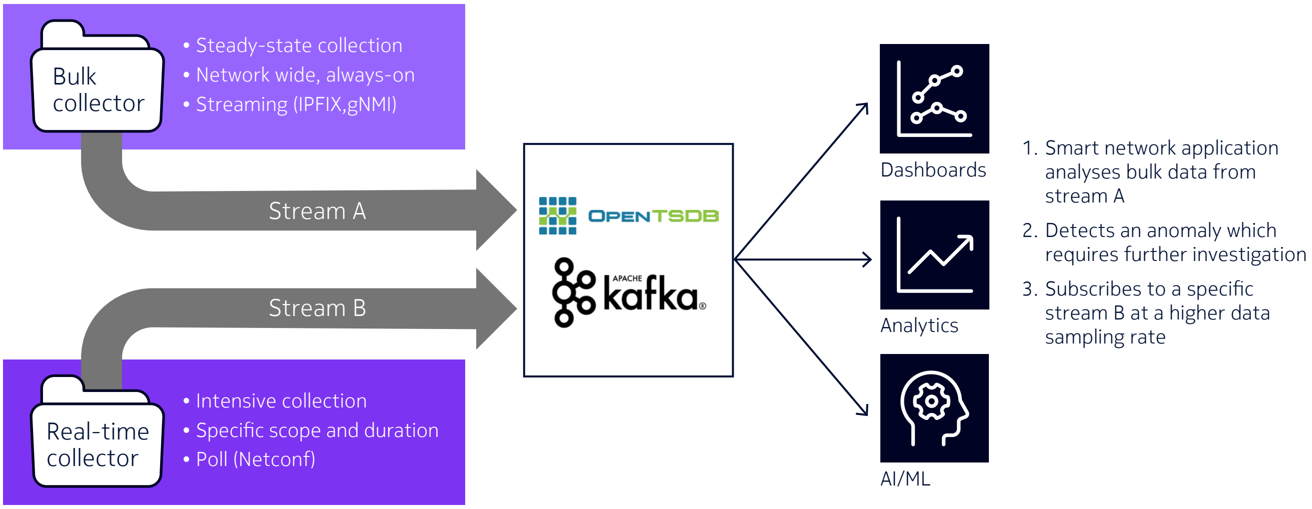
- **High precision**. Complex data in the right format – config, logs, alarms, counters.
- **Centralized**. Avoid a fragmented and incomplete view over distributed data.
- **Consistent**. High-quality data, real-time and historical, in normalized data sets.
- **Standardized**. Non-proprietary formats for direct comparison by ML algorithms.
- **Scalable**. Scalable database management in the cloud as requirements evolve.
- **Consumable**. Easy access for external systems without needing to duplicate data.
- **Customizable**. Operators can define their own models to process network data.
- **Low latency**. Time stamped and available for processing as soon as collected.
- **Secure and private**. Comply with Advanced Information Management (AIM) and General Data Protection Regulation (GDPR) requirements.

Modern data telemetry for efficient monitoring

Unlike traditional monitoring platforms, SDN programmable networks do not rely on continuously pulling data from the network elements in a rigid and pre-defined schedule. With push-based streaming telemetry, network elements will push data (stats, alarms, state, notifications, logs and other data) towards subscribed collectors based on a defined scope and frequency. Moreover, the model-driven approach offers the ability to subscribe only to data of interest, avoiding the transmission of huge volumes of unnecessary data. The programmability makes it possible to precisely define which data needs to be captured and processed by each application and consolidate subscription requests to the devices, which is important as monitoring each possible parameter in the network may not be practically feasible.

Applications may need to receive data streams at very short intervals. The telemetry framework needs to provide this capability without compromising the network performance. Therefore, subscription parameters should support runtime updates and re-configuration to dynamically adapt to evolving consumer application needs. Data collection could run in steady-state scanning (1-hour or 24-hour intervals) based on aggregated data sets for the entire network and, once an anomaly is detected or observed, the scanning could move to an intensive collection mode with a faster streaming rate (5-second or 30-second intervals) for selected narrow datasets, devices or objects. The two modes of operation should be able to go hand-in-hand, even with multiple parallel instances, on different datasets.

Figure 3 – Adaptive/interactive data telemetry



1. Smart network application analyses bulk data from stream A
2. Detects an anomaly which requires further investigation
3. Subscribes to a specific stream B at a higher data sampling rate

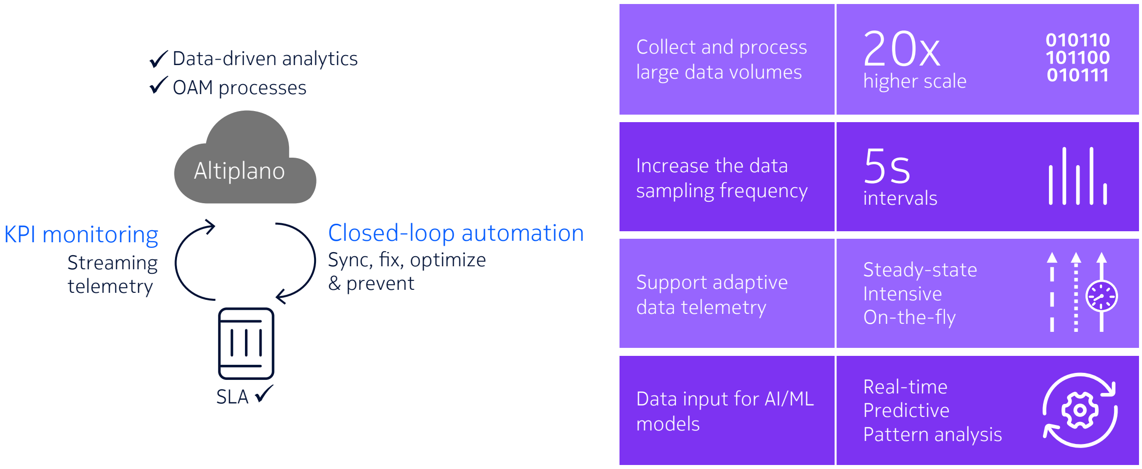
Apart from the usual performance counters of network resources, several other types of data need to be supported. Namely, key performance indicators (KPI) computed on top of performance counters providing abstract or aggregated information, alerts of abnormal conditions (faults), threshold crossing alarms (TCA), inventory, logs, traces and notifications of events generated by a system.

There are two data subscription modes available to make the streaming more efficient:

- **Periodic subscription** is perfect for cadence-based metrics that constantly change, such as traffic utilization, protocol statistics or hardware diagnostics.
- **On-change subscription** is suited for state information, which changes at more random intervals, such as service health, hardware state or admin up/down.

Customer use cases

Figure 4 – Efficient data telemetry sets the foundation for automation



Collect and process large data volumes	20x higher scale
Increase the data sampling frequency	5s intervals
Support adaptive data telemetry	Steady-state Intensive On-the-fly
Data input for AI/ML models	Real-time Predictive Pattern analysis

Typical service provider operations relying on high-quality data telemetry are:

- **Network monitoring** to understand how network and services are performing.
- **Network assurance** to perform preventive remediation of network anomalies.
- **Network engineering** to measure against design and performance targets.
- **Network planning** to provide design, upgrade, and dimensioning advice.
- **Customer care** to provide insights to resolve customer trouble tickets.
- **Marketing** to launch new products and improve returns on upsell campaigns.

We've identified four areas for SDN telemetry based on requests from our customers.

Handle large volumes of data

- A European operator wants to monitor the signal attenuation for more ONTs in their network. The degree of attenuation is an important consideration in the design of a PON network and requires information to be processed from both the transmitter and receiver sides. In disaggregated SDN architectures, it is possible to monitor 10x-20x more devices vs traditional networks.
- An Asian operator wants to filter the amount of alarm notifications that are reported to their OSS systems. With more than 6,000 nodes, the number of alarm notifications coming from ONTs can be huge in the cases of power outages or network upgrade procedures. With a more intelligent SDN controller, incoming alarm information can be processed, correlated and grouped per PON to optimize the information. Passive topology correlation can detect a feeder fiber or splitter issue when multiple ONTs on the same PON generate a loss-of-phy alarm.
- A North American operator reaches the limit of the number of User Network Interfaces (UNIs) that can be configured and monitored per OLT. In a disaggregated SDN architecture, the number of UNIs scales per line card, so it supports up to 16x more UNIs per OLT. On top, the bulk data collection created one big file, whose file size was reaching the limits of TFTP/SFTP file transfer protocols. SDN architectures use the IPFIX protocol, which is a more efficient method with individual caches allowing it to retain more metrics.

Increase the data sampling frequency

- An Asian operator wants to have uplink utilization and PON utilization data every 60 seconds to be able get a better view on bursts for their internal monitoring and predictive care applications. Another operator wants to do the same but for a set of individual priority ONTs. They also want to associate a threshold and send an alarm during threshold crossing alerts. With a traditional NMS they were limited to fixed intervals for all network parameters but, with streaming telemetry, the frequency can be increased for specific network parameters.
- A European operator wants to receive the PON utilization data at high frequency (every 5 seconds) to increase the accuracy of their traffic statistics. With traditional NMS network monitoring they are limited to 5-minute average values, and they cannot investigate congestion and detect packet discards that happen due to short traffic bursts.

Support adaptive data telemetry

- A North American operator has a set of different OSS applications, each wanting to consume a different data set, applicable for high-frequency collections at a network-wide scale. With SNMP, all applications would collect basically the same level of data without any differentiation in requirements. SDN controllers can have multiple IPFIX exporters for different data sets and allow NBI applications to ingress data from specific Kafka topics assigned for those data sets.
- A European operator monitors the top 5 PONs with the highest utilization every 15 minutes and would like to do the same for packet drops or any other user-defined metric. This is configurable in a SDN controller, but the success will depend on the required metric and scale. On-the-fly calculations would be possible if based on the currently received data set. If data from previous intervals needs to be fetched every 15 minutes, this may not be possible at the scale of millions of ONTs.
- An Asian operator wants to perform once-in-a-day scans of the G.fast performance for the entire network to detect anomalies that may impact the bitrate and stability of the lines. For an identified subset, they want to trigger live collection to monitor operational data more frequently (every 5 seconds) for a duration of 10 minutes. All selected (“suspicious”) nodes need to be polled daily. The SDN controller can dynamically adapt to these consumer application needs.
- A European operator wants to monitor new subscribers closely for a few days to make sure the service experience is smooth. Key optical and traffic metrics should be monitored at high frequency for only the new subscribers. Identification of signal degradation or dropped packets should be flagged immediately. After three days of monitoring, new subscribers can be released from rigorous watch.

Data input for AI/ML models

- A Middle Eastern operator wants to monitor closely a set of ONTs from priority subscribers, evaluating the performance based on criteria, such as traffic utilization and optical performance. They also want to run automated tests and troubleshooting workflows when problems are detected. Network care tools can than extract the symptoms, map them against known issues database, and suggest recommendations based on existing resolution actions for the historical incidents.
- A European operator wants to collect network data and run a multivariate analysis to identify anomalies in the network. Smart data tools and machine learning allow network data to be clustered and then identify multi-dimensional outliers without set thresholds. For example, they can evaluate SFP performance from different suppliers in the entire network on a range of parameters and identify problematic SFPs before they cause service outages.
- A Middle Eastern operator wants to train a machine learning model which represents their broadband network to improve network planning decisions. The network modelling needs the dataset as accurate as possible to reflect the network characteristics. High-frequency streaming telemetry (at 5 second intervals) is a key requirement for building such a data model. Broadband usage includes several burst patterns (e.g. speed tests, downloads) which are not visible in low-frequency metrics.
- An Oceanic operator wants to improve the overall performance of their PON network and ensure fair peak rate availability to all subscribers. Long downloads are becoming more frequent and block other users from using Gigabit peak rates. By closely measuring traffic utilization, they can apply smart congestion management in the OLT (dynamically optimizing scheduler weights and shaper rates) and improve the subscriber noticeable peak performance for users during busy hours.

Conclusion

Data and insights are essential for efficient network operations. The arrival of SDN in the broadband access network enables operators to make far more effective use of data, especially when it comes to automation. Key to all this is telemetry, specifically intelligent and flexible push-based streaming telemetry where network operators can pick and choose the amount, complexity, and frequency of the data they analyze. Much of this can now be in real-time, while SDN-enabled solutions are capable of monitoring 10x-20x more counters compared to legacy NMS solutions. Through good data governance, telemetry coupled with SDN enables many operator use cases across network monitoring, assurance, engineering, and planning, through to customer care and even marketing opportunities.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: (July) CID213261