The Fiber Folding Ruler

Creating a Common KPI Language for Operating Fiber Networks

A technical paper prepared for presentation at SCTE TechExpo24

Robert-Jan van Minnen, MSc.
Senior Manager Network Performance & Analytics
Liberty Global – Liberty Tech – Fixed Networks
rvminnen@libertyglobal.com

Jason Rupe, Ph.D.
Distinguished Technologist
CableLabs - Technology - Wired
j.rupe@cablelabs.com

Table of Contents

Title **Page Number**

List of Figures

List of Tables

1. Introduction

Cable and fiber networks are evolving and becoming increasingly hybrid. The same applies to network- and service organizations. Network operations engineering and staff can be burdened by the expansion of the technology they must operate, and the differences in tools to do so. This condition calls for an evolution of the way performance is reported in a unified way. In turn, this requires the telemetry from networks to provide sufficient and comparable data which feed into business processes. The inspiration is that having a fresh look at the combination of hybrid fiber-coaxial (HFC), fiber and Ethernet telemetry, a new and better framework can emerge. It is also an opportunity to reset the panes so that upcoming technologies such as Artificial Intelligence can develop on a solid database.

This paper explores and captures both the challenges and opportunities to get the best of both worlds. By peeling down the essence of existing key performance indicators (KPIs) it becomes obvious there is in fact a lot of similarity. But also new territory is entered with the management of Converged Interconnect Networks.

The creation of a common language and framework with respect to network KPIs is proposed. The process of developing this common language and framework is powered by an expert CableLabs working group (the optical operations and maintenance (OOM) working group). This paper strives to galvanize a larger audience into action in support of this quest and to stimulate an increased joined collaboration and co-development. In support of this goal, building blocks are provided.

2. Scope

This initiative originates from the evolving cable industry which is proficient in shifting boundaries to keep meeting customer demands. Part of this evolution is a growing fiber richness, in some cases up to full fiber to the home (FTTH). While being ready to adopt already existing knowledge from operating cable and fiber networks, sometimes a step back is taken to create a wider view and see if better solutions can be found. To maximize synergy between initiatives, the development described in this paper has been added as a workstream to the CableLabs OOM working group with the following mission and vision.

2.1. Mission

<u>Drive the creation of a common 'KPI language' for operating networks</u>.

Develop the framework for tools to integrate and identify ways to combine KPIs around decisions and actions that are common.

2.2. Vision

<u>A commonly accepted standard for </u>KPIs that relate telemetry data from networks to value for customers and operators. This will be use-case informed, and oriented around the needs of network operations efficiencies.

3. Approach

This chapter describes the logical steps in the paper which are detailed out in the following chapters.

3.1. Setting the Target

The investigation starts with identifying the business processes that are in scope. In other words, which business purposes or use cases must be supported with the KPIs.

Working from the business values of cost, performance, and customer experience of the use cases, the desired KPIs are defined. These are validated according to the required properties of a good KPI.

3.2. Discover What is Available

A comparison is made with common practice KPIs for HFC. Similarities will support smooth transition from one technology to the other and operational benchmarking.

Existing network Element Management Systems (EMS) provide pre-processed telemetry. Though often vendor specific, these may give practical guidance towards required KPIs.

Because fiber networks are all around, operators will have developed best practices including KPIs. Though these may be operator specific, these will also provide practical guidance towards required KPIs.

3.3. Creating the Map

Working from the use cases, map out:

- Direct fit
- Near fit – adjust or new development?
- New development
- Priorities related to business needs
- Obsolescence

3.4. Fulfillment

Following the priorities in the map:

- How to obtain telemetry
- Rules for aggregating
- Logic to define thresholds in relation to business value

3.5. Recommendations:

- KPI map
- Implementation roadmap
- Standardization
- Further development

4. Setting the Target

4.1. What is a Good KPI?

A good KPI, helps in achieving business objectives. To create focus, the following required properties for each KPI are proposed:

- Clear:
  The definition and process of data collection and aggregation should be easy to understand.

- Relevant:
  KPIs should directly or indirectly relate to the cost of operating a network, service performance or customer experience.

- Comparable:
  KPIs from different networks should be comparable. This implies independence of equipment brand and market structure.

Some resources suggest that a KPI should be Specific, Measurable, Attainable, Relevant, and Time-Bound (SMART). While these are good criteria for a KPI, they are not always appropriate. For example, a KPI of 100% availability is not achievable, but may still be a good target to work toward and may be achievable for periods of time. Further, a focus on SMART as the criteria misses several important additional features: a linear KPI is easy to understand how close to a target value the current performance is; a KPI causally related to a desired feature to control and to the levers the business has for controlling them is very important; a KPI that everyone can understand is more likely to be achieved.

Additional criteria to be considered, taken from [4]:

- “Can it be easily quantified?
- Are we able to influence/drive change using this KPI, or is it out of our control?
- Does this KPI connect to our objective as well as overall strategy?
- Is it simple to define and understand?
- Can it be measured in both a timely and accurate manner?
- Does it contribute to a broad range of perspectives – i.e. Customer, Financial, Internal Processes, Learning and Growth?
- Will it still be relevant in the future?”

4.2. Aggregation

KPIs are typically aggregated numbers. Depending on the KPI, the aggregation includes:

a. Timing: the timeframe that is reflected and time-related rules (e.g. the busiest hour)
b. Summing rules: to preserve consistency when adding data from network parts together
c. Telemetry data: defining the exact data points in use
d. Process data: inputs besides telemetry data (e.g. truck roll count)

4.3. Telemetry

Telemetry is an essential element of a practical KPI framework. While the generic definition assumes a telecommunications network for transporting the measurements, in our case it is the network itself that does the transport.

Definition: Telemetry is the in-situ collection of measurements or other data at remote points and their automatic transmission to receiving equipment (telecommunication) for monitoring.[4]

Benefits of the network probing itself are:

i. No need for a separate transport network
ii. Integrated paths for transport
iii. Controllable probing cadence and volume

Challenges:

iv. Network elements are not measurement devices. Results include certain error margins which must be considered when used. E.g. reported power level is +/- 1dB.
v. Dependent on certain network layers to operate. A failing link may conceal or alter underlying data.

Note: some definitions separate alarm data from measurements. In the scope of this paper however, these are both captured under the definition of telemetry.

5. Business Processes

The business processes to support are considered in a logical order to ensure completeness and transparency.

1. Acceptance from construction
2. Provisioning
3. Installation
4. Maintenance
5. Fault management
6. Performance management
7. Capacity management

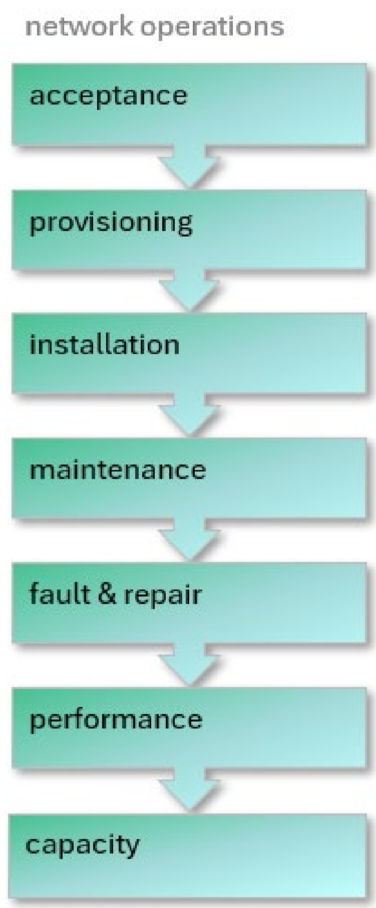


Figure 1 – Network Operations processes

5.1. Acceptance from Construction

During construction of a fiber to the home (FTTH) network a separation between activities and responsibilities can occur:

- OLT infrastructure. This is the network up until the optical handover from OLT's (blue box).
- PON infrastructure. The passive optical network from the OLT handover up until the point where the customer connection (drop) can be made (orange box).
- Optical drop. The connection from the last splitter, tap or connection box to the ONU in the premise (green box).

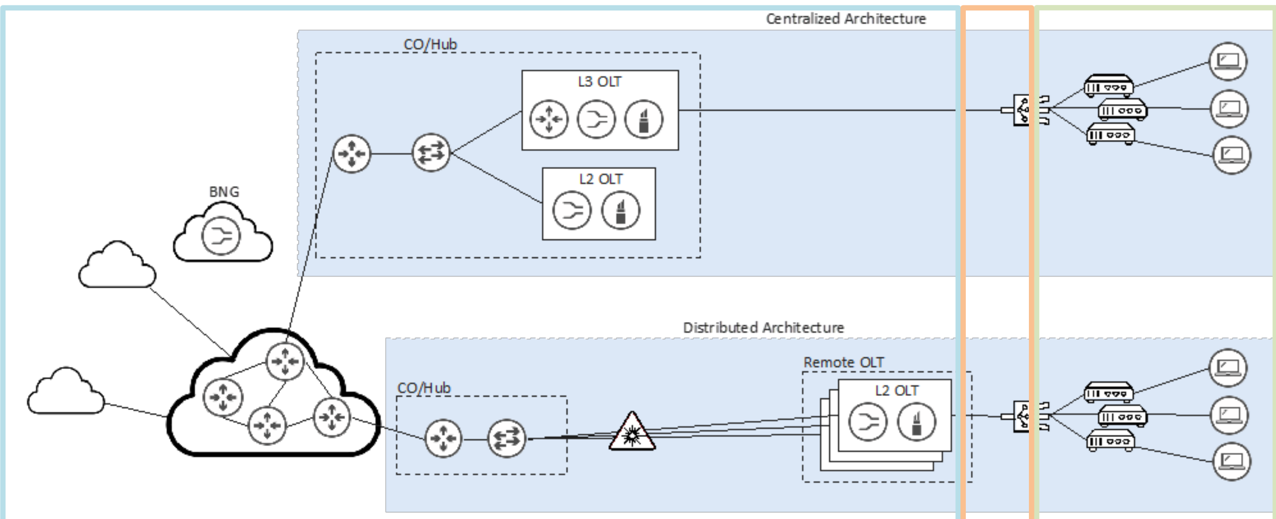


Figure 2 - Word Dialog Box Shown When Updating Fields

To accept these slices into operation, the built network must be certified with a so called 'birth certificate'. The separation requires definition of three birth certificate types to prove presence, compliance with quality standard and to store initial values. These initial values will become essential in later phases for trouble shooting and when changes are made.

In a later phase of service installation, a fourth type is required – installation birth certificate-.

Table 1 – Birth Certificates concept

Certificates	OLT	PON	Drop	Install
Presence	OLT-ID	OLT port ID, location	PON-ID, Address	ONU-ID
Quality	checklist	OTDR report	OTDR report	RX, tests*
Initial values	transmit power	return loss, distance	return loss, distance	RX, TX, bias, test results

* Installation tests are dependent on the requirements for the service and operation and are not included.

<u>A KPI to monitor acceptance for construction would be based on the number of built elements (OLT, PON, Drop, Install) and the numbers that pass the quality criteria</u>.

The initial values would be stored for comparison and trend analysis in later phases.

5.2. Provisioning

Like construction, provisioning is about providing the right settings and configuration to network elements to perform their intended role ('telecommand' 2]). Though the actual provisioning falls outside the scope of this paper, a KPI can be defined as the actual provisioned state in comparison to what is intended. Using telemetry this state can be read and compared with an external source with intended settings.

Table 2 – Provisioning state

Provisioning state	OLT	ONU
Read	settings	settings
Intended	settings	settings
Match	0 / 1	0 / 1

A KPI to monitor the provisioning state would be based on the number of elements (OLT, ONU) and the numbers that have a matching provisioning state.

The provisioning state is obtained through telemetry. The telemetry read activity will be re-used in later phases such as fault management. While the KPI can be obtained in bulk and during quiet times, for fault management it should be individual and immediate.

5.3. Installation

This section focuses on activating a service for a customer. Before commencing this, the previous steps should be confirmed (OLT, PON, Drop, provisioning). But, in practice the process may combine steps.

For example, the drop construction is combined with ONU installation. The birth certificate of the drop construction could be skipped as nearly the same data can be obtained through the ONU telemetry. This is an acceptable cost saving but reduces the accuracy for the maintenance and fault management phases because the ONU measurement includes both its own deviation and the drop attenuation. When the ONU is replaced, the apparent drop attenuation may change. This must be considered in these phases.

Another example of combination is where an installed ONU is blank, reports itself (ONU-ID) and discovers its position (OLT, port ID) and obtains provisioning. This is a flexible and efficient process but requires complex support systems. Availability of these systems is critical for the installation process. Especially when customers install themselves because they have no means of verification.

Finally, CPE may be installed with services including Wi-Fi. All these offer telemetry elements which vary with vendor, service and operator.

A KPI to monitor the installation process would be based on the number of installations (technician or customer) and the numbers that passed the criteria for ONU Install and all intended services.

For management of the process, it is recommended to detail the results of failed installation in a structured way. See the relevant section in Fault Management.

5.4. Maintenance

According to a definition, maintenance consists of '*various cost-effective practices to keep equipment operational; these activities occur either before or after a failure*'. In an operator's practice, it is intended to minimize the occurrence of failures and impact on customers in the most economical way.

In a simple view, the order of things for an ONU:

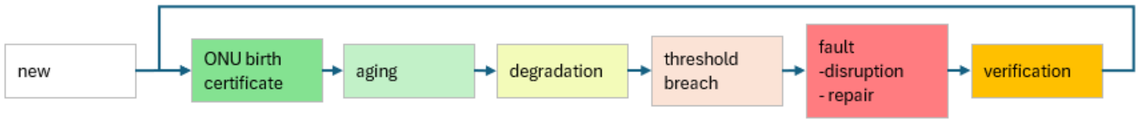


Figure 3 – simple maintenance flow

Typically, components degrade before they break. In many cases, this degradation takes time. When the degradation is detectable through telemetry and predictable with learned degradation-curves, there is the opportunity to correct before the break happens. This has big benefits in terms of costs and customer experience.

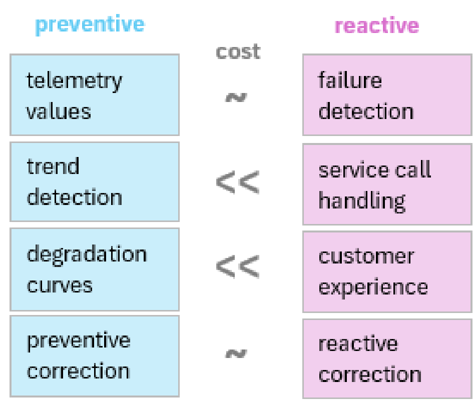


Figure 4 – preventive versus corrective maintenance

While it seems obvious to focus on prevention, there are challenges:

- The right telemetry must be available to detect trends
- A procedure or system must apply thresholds from learned degradation curves
- An optimization process must be implemented (correct just before it breaks)
- The benefits must be proven (like the fire alarm system: how to prove the business case)

These considerations provide guidance to the KPIs required. Note that these only reflect a high-level perception of the underlying process and telemetry.

KPIs to monitor preventive maintenance would be:

- CIN link laser bias
- OLT laser bias
- ONU TX, RX
- ONU X-GEM/bit errors
- # preventive corrections

For management of the process, it is recommended to detail the data about preventive corrections in a structured way like the way faults are registered (see: fault management).

The KPIs for reactive maintenance are part of fault management.

Because degradation curves and valuation of customer experience may only be indirectly measurable, there is a risk of not performing the right number of preventive corrections. Artificial Intelligence (AI) will become a useful tool to refine the curves and customer experience valuation. Standardization and transparency for AI-applications should be considered when KPIs and telemetry data are defined.

5.5. Fault Management

A fault is basically anything that doesn't operate within acceptable parameters. In technical terms: something is out of spec or broken. Important faults are those that link to a potential failure. If not for a resiliency mechanism in the system, the fault may impact service and lead to failure to meet intent. A fault, therefore, can trigger maintenance that is proactive, as opposed to a failure that triggers a reactive repair to restore service. Because a fault is usually prevented from becoming a failure due to some resiliency mechanism, it can be managed by capacity of that resiliency. For example, an impaired fiber can be compensated by increasing the Tx levels of the laser, but only up to a point.

This definition implies that for every telemetry data point, boundaries should be set to determine in or out of spec. It is likely that many of those data will also be used for maintenance to prevent values to become out of spec. However, even if within specifications, there may be system behaviors that are indicated in telemetry that are worthy of attention. For example, temperature may remain within specifications but vary more widely than expected. Range or variance statistics on the temperature data may suggest problems with fans, filters, or other problems that should be addressed proactively. Also, temperature fluctuation may lead to shorter useful lifetime of hardware.

In essence, faults are detected as specific cases of the data used for maintenance. But besides this telemetry data, faults are reported in by customers (service calls) and the Network Operations Center (NOC).

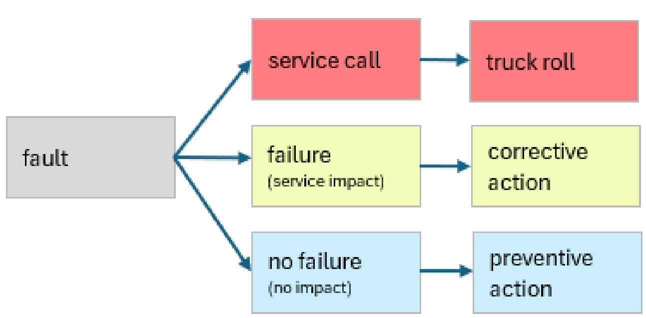


Figure 5 – Example faults leading to action

To combine these data, it is strongly recommended to register faults in a structured way. As an example, the ‘fault categorization’ as in appendix A can be used. This is a practical example that follows the general architecture of the network. It is simple enough to be used in the field but detailed enough to analyze and drive improvement programs.

The frame in which faults are captured must be coherent with the general architecture and related telemetry. In essence it is a two-dimensional table with the third dimension being time.

A KPI to manage faults would be the total number of faults per period divided by the number of elements in each cross section or in any other required ratio.

5.6. Performance Management

Performance in the context of this paper relates to the achievement of qualitative technical parameters from services from the perspective of the business.

The most prominent perspective is of course the perception by the customers. Market and regulatory evaluation through open tools like SamKnows, Umlaut etc., should also be considered. Finally, business decisions about technology also rely on information about technical performance of network elements such as fault rates.

Performance can generally be described in the following dimensions:

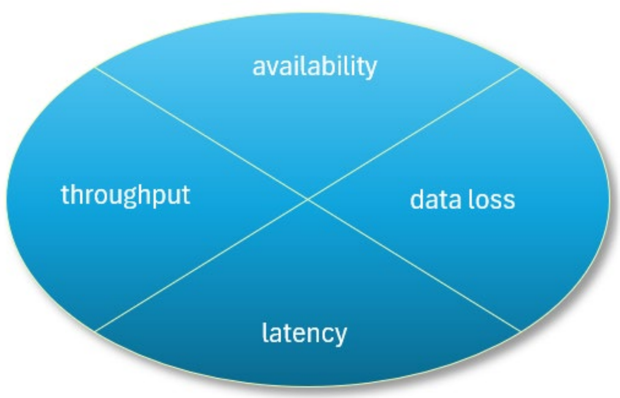


Figure 6 – dimensions of performance

5.6.1. Availability

Also referred to as uptime, this is a combination fault rates and time to recovery. In a simple formula:

availability = 100% - failure events in the time interval x time to recovery

While fault rates are discussed in the chapter about fault management, the time to recover needs addition of timestamps. For example, the time of changes in element state.

The availability of a service requires all elements that form the service to be available. If the availability data from all elements forming the service is known, the service availability can be calculated. In many cases however, it is more efficient to use telemetry from a higher layer of the connection or the from the service itself to determine its availability.

A KPI for availability would be based on the failure rates and recovery times, measured as close as possible to the service in scope. See [5] for more guidance.

Note that target availability figures may vary with time of day or week, for example to allow for service windows. This implies that multiple versions (calculations) of the KPI could be needed.

5.6.2. Throughput

In essence this is the amount of data or payload per unit of time. Also referred to as traffic or bandwidth. The typical notation is in megabits per second (Mbps) but also kbps (per customer) and Tbps (per network) are used. Note that these are speeds rather than total consumption but since the unit of time is always a second, conversion is simple. In parallel to what is commonly used in DOCSIS® specifications, it is recommended to use 'Downstream' and 'Upstream' instead of 'in or RX' and 'out or TX'. This avoids ambiguity.

The amount of data is typically determined by comparing bit-counters between two timestamps. Samples from these counters can be taken at various intervals. For example, every 5 minutes, 15 minutes, hour etc. The resulting throughput is therefore the average during the sample time. It is important to realize that between two samples, the traffic varies. This is known as the bursty or statistical behavior of traffic. In the SCTE paper from 2022 'the speed triangle' the statistical relation is discussed.

Typical example of traffic KPIs:

a. during evening peak hours (e.g. 8-12pm)
b. during the busiest hour of the week
c. at busiest 15-minute sample
d. for one month

A network exists based on sharing media. In those media, traffic is added (or subtracted). Similarly, within the same sample, the traffic KPI's can be added/subtracted. But if the timestamps are variable as in 'busiest', the calculations need a statistical approach. The logical approach is to take traffic samples at different points in the network:

- ONU: downstream and upstream per hour
- PON: downstream and upstream per 15 minutes
- OLT uplink: downstream and upstream per 15 minutes
- CIN links: downstream and upstream per 5 minutes

The KPIs derived from these could be downstream and upstream traffic:

1. ONU at the busiest hour of the PON
2. PON at the busiest hour and busiest sample
3. OLT uplink at the busiest hour, busiest sample and week average

4. CIN links: at the busiest hour, busiest sample and week average

Note that many of these KPI can be normalized for comparison. For example, a link capacity can be 10 Gbps, and utilization may be 5 Gbps, leading to a consumption of 0.5. This compares to a 1 Gbps link with utilization at 0.5 Gbps having the same consumption proportion.

5.6.3. Data Loss

During transport, some data do not reach the destination. For a transport network this is undesired and so it must be monitored. The following potential reasons for data being lost or dropped are recognized:

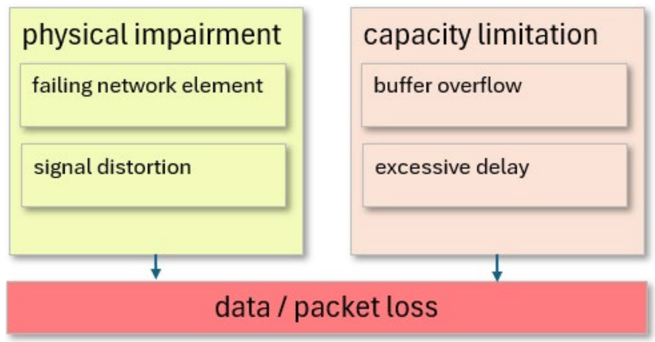


Figure 7 – reasons for data being lost

5.6.3.1. Physical Impairment

On the lowest level, networks are designed to cope with individual bits being erroneous. On this level, a bit error equals the loss of a bit. Bits are grouped into codewords of typically 16 to 256 bits to allow forward error correction (FEC) mechanisms to correct the error. If the number of errors is below a certain threshold, no data is lost because of protect bits.

When the number of errors is too high, the FEC cannot correct and the codeword is lost. Because degradation can progress and lead to increasing errors that grow, it is good to track the correctable codeword errors as an indicator.

In G-PON/XGS-PON, codewords are put together in frames of variable length with the XGS-/G-PON Encapsulation Method (X-GEM). Through a header error check (HEC), errors of these frames can be detected and corrected to an extent.

5.6.3.2. Capacity Limitation

When during a certain interval, more data is offered than the medium can transport, it is rejected. When a preceding buffer is available it can be stored in a queue. When there is no buffer available or when the buffer is full, there is no other option than to neglect the data until there is room.

Some data must be delivered in the right sequence within a certain time, for example the voice in a phone call. If some packets suffer too much delay -for example through multiple buffering- they become useless and can be deleted before delivery.

KPIs to manage data loss could be errored X-GEM packets and dropped packets as a percentage of total offered (or transported) packets on CIN, OLT and ONU's. One may also want to collect buffer utilization statistics.

5.6.4. _Latency

Latency in this context reflects the time it takes traffic to traverse the network and its variation. In a simple view, the delay is caused by a few components as depicted below.

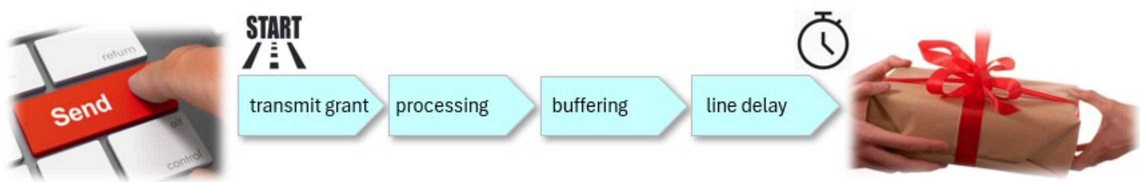


Figure 8 – latency components

1. A scheduler grants 'airtime' to transmit. The availability of airtime depends on how busy the medium is and how the scheduling is programmed. This means that the momentary utilization of the medium has a direct influence on the latency of data.
2. Processing time of equipment is constant as airtime has been scheduled. However, it gives some statistical variation depending on timing of the signals.
3. Buffering delay depends on the size and structure of the buffer. Under nominal conditions it should be relatively short and constant. But as it depends on the momentary utilization, it will vary with this. When the buffer is full, delay times can become excessive of packets are dropped.
4. Line delay is typically constant as it depends on the physical distance.

Measuring latency directly for an entire network seems not practical as it would require adding timestamps to packets and a collection system. It would add additional traffic and with that influence the results. Instead, collecting samples through probes in for example CPE can be used. These give indications about the network latency in a structured way. It must be considered though, that sporadic excesses from buffer overflows or parallel traffic on the CPE may be invisible in the results.

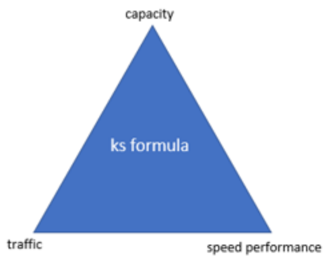
Speed tests can be considered a specific case of latency test; the time it takes a large test file to traverse the network. Since the same resources are needed to automatically perform speed tests, it is recommended to use the opportunity and combine these with latency.

A KPI to monitor latency would be the result of a structured latency- and speed test by a representative number of probes.

5.7. Capacity Management

The aim of capacity management is to ensure that the right amount of capacity is available. This is always a balance between the ability to deliver the required level of service and cost.

The minimum amount of capacity needed is basically a function of the traffic or throughput and the required speed performance. See also the section throughput where reference is made to the speed triangle study on this relation. The method can be used to drive capacity growth based on traffic predictions and desired speed performance. It is technology-independent and can be calibrated using empirical data, for example from probes performing real life tests.



capacity = utilization * s + speed * k

s = service level factor, typically 1
determines the chance of successful speedtest

k = desired speedtest result, typically 1
determines the speed achieved

Figure 9 – the speed triangle

This method can be used on basically any part of the network and can be converted to provide practical thresholds for upgrading. It can also be used to estimate the available speed room and chance of successful speed tests in a live network using common capacity and traffic data.

The option to estimate speed room also enables a new approach for the concept of congestion. The traditional approach for congestion is to keep utilization (i.e. traffic as percentage of capacity) below a certain threshold (e.g. 60% during evening hours). This is a coarse method to avoid congestion which is typically defined as the utilization reaching a high level (e.g. 90%) during certain hours (e.g. the busiest hour of the week). This method is a coarse approach which gives limitations because it doesn't include specific speed requirements. These limitations will increase when on a fiber network multiple service types with different requirements are combined.

The speed triangle gives the ability to define congestion according to general traffic theory which defines three phases.

1) Free flow of traffic (speeds can be achieved as planned)
2) Slowdown of traffic (speeds may not be achieved at times, but still flowing)
3) Traffic jams (speeds can drop to zero, packets may be lost)
   The traditional congestion threshold aims to avoid this phase

The speed triangle will be expanded to include multiple services (e.g. with committed information rate) in the future, but this already implies additional requirements for the traffic data from different services present on the network.

Besides the capacity on transport layers, the capacity of processors or storage within equipment may need to be managed. This is not included here but may need additional KPIs, based on properties of network elements (BNG, CIN, OLT).

KPIs to monitor capacity management would include the available capacity on the layers of the network (CIN, OLT, PON) and the minima required because of the traffic and speed requirement. With the same mechanism, the probable speed performance can be predicted. In addition, the amount of congestion should be measured in terms of phase 2, slow down.

Note that resiliency mechanisms, and most any network resource that is virtual, physical, or otherwise quantifiable, can be managed like capacity. For example, even the telemetry delivery of a network device will have limited capacity and can be managed as such. Refer also to the earlier example of TX power level.

6. Comparisons with HFC

6.1. Comparing KPIs

On the high level of the KPIs above, a comparison can be made with common practice with HFC networks. The simplified table below shows a high degree of similarity.

Table 3 – KPI comparison FTTH - HFC

FTTH	HFC	note
built elements passing criteria	existing though different elements	include total installed base
elements with provisioning state	existing though different elements	reflects total installed base
installation passing criteria	existing	
maintenance KPI's bias/RX/TX/errors	existing though different elements	requires telemetry
preventive corrections	existing	integrate in one fault categorization model
fault rates, truck rolls	existing though different elements	integrate in one fault categorization model
recovery times to calculate availability	existing	
traffic at CIN, OLT, PON, ONU	existing though different elements	requires telemetry
errored and dropped packets at CIN, OLT, ONU	existing though different elements	requires telemetry
latency and speed tests	existing	
capacity available and minimum required	existing though different elements	prepare for combined services
congestion	existing though different elements	add new congestion measure

The following preliminary conclusions are suggested:

a) All KPIs basically exist with HFC.
b) Most KPIs are based on data from sources different from HFC.
c) It is logical to capture data in a structured model to support management.
d) The model could be structured according to a standardized network architecture.
e) The HFC model could be the blueprint, smoothening transition analyses.
f) Requirements for telemetry should be aggregated because some occur multiple times.

6.2. Alignment

Limits specific to a technology can be translated into limits of general qualities that the technology provides. For example, the amount of capacity that can be added to a DOCSIS network to serve a particular cable modem is limited by that cable modem's ability to transmit more power for the additional channels. For most optical access networks, adding channels is not an option, so there is no direct parallel for PON.

But there are parallels between DOCSIS and PON. For a DOCSIS network there is a limit in the amount of power that a cable modem can transmit in a channel to overcome impairments and limits in the coax network, and likewise in PON networks, the amount of power that the ONT may transmit is limited and therefore the amount of headroom for ONT transmission can be determined. Because these two can be normalized to between 0 and 1, a performance measurement that equates the two could be created, with parallel meanings.

From a service perspective, access network service has limited dimensions: throughput or bandwidth or bitrate, latency or delay, data (bit, packet) loss, jitter or delay variation, and uptime or availability as a combination of time to fault or failure and time to recovery or repair. KPIs that focus on the support of these dimensions are important. To be more complete however, additional dimensions are needed. Delay is not the only dimension that can vary, and not the only one which impacts service when it varies. Bandwidth variability and other performance measures that vary will impact service too, and reliability and availability problems are but extreme variability in performance. Tracking the first five dimensions and their variability over time therefore covers access network service delivery well, and KPIs that address those dimensions can assure operations are aligned to service. And because these dimensions are technology agnostic, KPIs can be unified across technologies, including DOCSIS and PON.

Any factor that limits one of these dimensions becomes a limit on the service. Without sufficient capacity on the network, no other customers can be added. A fault in the network that leads to data loss cannot be tolerated indefinitely or for all use cases. Excessive downtime leads to customer dissatisfaction and unfulfilled guarantees or service goals.

Network operators monitor component and system state on the network to assure functionality. They also monitor network state to find changes in network behaviour that indicate faults and failures. Operators monitor relevant indicators relating to repair to assure downtime is minimal, and resources are used well. They monitor provisioning and service state to assure service is established and assured fully. Among these categories, many network operations KPIs are defined.

From a network performance perspective, there are measures of performance that indicate limitations in resources that support service. Capacity management is not just for network bandwidth capacity. It applies to any resources in the network that is limited but needed to provide service including any resiliency mechanism. For example, DOCSIS has profile management which enables the trading of bit loading (throughput) to improve service reliability, cyclic prefix and roll off to trade time (throughput) to improve transmission reliability (resiliency against echoes and burst noise and ingress), and equalization to apply limited power to provide the highest possible bit loading and dynamic range in the system.

For a PON system, which mostly transmits in limited frequencies, has fewer resilience mechanisms.

6.3. Differences

6.3.1. Multi-channel versus single channel

DOCSIS bonds multiple channels in the HFC spectrum to increase bandwidth. The complexity of managing the usage and performance of these channels simultaneously does not exist in current FTTH applications. This eases especially the gathering/aggregation of data and procedures within fault management.

6.3.2. Steeper cliff

DOCSIS has developed to be very robust against signal impairments. This means that signals and impairments can vary over a wide range with only minor data distortion until the connection is lost. Observations are that with fiber, this transition from good to bad is much steeper when optical signals are impaired. The upside of this is that thresholds -for example for optical signal levels- can be set uniformly. On the other hand, it reduces the time available for maintenance procedures to react upon signs of degradation. This should be mitigated by the quality of Maintenance.

6.3.3. Signal distortion

The typical issues in a coax network caused by unwanted interference (ingress), radiation (egress) and non-linear connections (common path distortion) do not seem to play a role in a fiber network.

In DOCSIS, we have RxMER, spectrum analysis, channel estimation, and pre-equalization, because we use wide frequency bands that can be impacted by physical impairments. To manage these impairments, we have tests and queries from network devices that we can use to identify and localize faults before they become failures (PNM). In PON, this is not the case. In PON, we rely on power levels that are queried. But that is not useful for locating issues; an OTDR is therefore needed. But even that OTDR is usually limited in the frequencies it can analyze. Perhaps we need a sweep-like function for PON, to make maintenance proactive for optical networks too.

On a fiber network other types of distortion occur. The distortions that can appear as degradation should be detected with telemetry as much as possible and propagate to the maintenance process as discussed.

6.3.4. Service group size

Access networks work from the principle of aggregating the traffic of several customers on central equipment. In DOCSIS we would call this a CMTS service group, for example connecting 400 homes. On a PON network it would be an OLT port with 64 homes. Assuming a constant penetration with customers and speed performance, this is a reduction by a factor of six. The effect is a significant reduction in efficiency in terms of traffic/capacity on the PON (factor 4-5 in this example).

The traffic is further aggregated in the OLT which increases efficiency again to similar levels as in a CMTS. The speed triangle provides tools to model the network there as efficiently as possible. It is recommended to apply the tools on the OLT port level as well because the sensitivity for single user behavior has also increased significantly. This sensitivity increases further if services with different quality of service are on the same PON.

6.3.5. CIN

With FTTH, the CIN network is an intrinsic part of the access network. This implies that processes and thus KPIs apply there in the same way. Because it is a converged network, it can also transport other services – for example mobile backhaul or business services. These services may have different requirements than FTTH. Care must be taken that KPIs and telemetry are compatible with any possible service considering that only thresholds can be different. For example, a service with a committed information rate (CIR) will impact the available capacity for residential services with a statistically defined speed room. Further work is recommended to verify the impact of coexisting services with additional requirements. This also applies to the OLT and PON level as other services can coexist there as well.

6.4. Scanning the market

A first and very limited survey across the market did not yield much complementary information. Focus seems to be at building fast and efficiently. Maintenance is limited because of relatively young networks and fault management is not much different from cable networks. The perception is that the capacity of the fiber is more than enough which limits capacity management to only managing the number of customers per OLT port.

While this seems disappointing – we may just not have found the right information – the fact that the industry is able to produce these insights and finds ways to obtain the right data gives an advantage in optimizing the performance and economics of a fiber network.

7. Fulfillment

KPIs as described are typically aggregates from metrics obtained from the network through telemetry. On top of that, data from other sources such as work force management can be added.

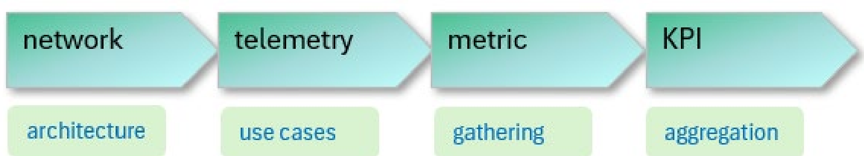


Figure 10 – aggregation of KPIs

7.1. Network

The network is considered in the way it is designed. As mentioned, it is recommended to capture telemetry in a universal way in which basically any network architecture fits. This implies simplifications but it enables structured collection of telemetry and other data which is required to unify the resulting tools and KPIs. This unified architecture forms the basis for data collection, metrics and KPIs.

7.2. Telemetry

Use cases as mentioned are specific applications (for example fault finding) where telemetry plays a crucial role. In fact, while definition of the telemetry of the uses cases has not been finalized, the KPIs in this paper can be related to these:

Table 4 – KPI to use case category

FTTH	use case category
built elements passing criteria	NOC
elements with provisioning state	Provisioning
installation passing criteria	Birth certificate

maintenance KPI's bias/RX/TX/errors	Link Health
preventive corrections	Fault management
fault rates, truck rolls	Fault management
recovery times to calculate availability	Fault management
traffic at CIN, OLT, PON, ONU	other
errored and dropped packets at CIN, OLT, ONU	other
latency tests	other
capacity available and minimum required	other
congestion	other

7.3. Metrics

Telemetry needs to be gathered in a system. It is recommended that the system:

- Collects data with the fastest cadence as advised by either the use cases or the KPIs
- Adheres to the proposed unified architecture format
- Pre-processes data to reduce volume and eliminate obvious errors respectful of all current and future use cases
- Retains a full data set for short periods of time to enable currently unforeseen analyses

7.4. KPIs

To aggregate metrics into KPIs, rules are required. For some it would be a simple addition while for others specific samples must be selected and/or combinations must be calculated. It is important that the algorithm of aggregation is clear and preferably universal. A different aggregation can give different KPI results which would render comparisons, benchmarking and service level agreement (SLA) purposes extremely difficult. A full definition is beyond the scope of this paper, but some recommendations can be given.

Table 5 – KPI aggregation notes

FTTH	KPI aggregation notes
built elements passing criteria	counting + adding to installed base
elements with provisioning state	counting + % of installed base
installation passing criteria	counting + comparison with work orders
maintenance KPI's bias/RX/TX/errors	apply thresholds + % of deviations of installed base
preventive corrections	counting + % of installed base
fault rates, truck rolls	counting + % of installed base
recovery times to calculate availability	compare timestamps from fault management + workforce
traffic at CIN, OLT, PON, ONU	busiest hour, busiest sample and week average
	busiest = on the unit sample + if needed on the full network
errored and dropped packets at CIN, OLT, ONU	% of offered traffic on elements + service layer
	apply thresholds with % of deviations of installed base
latency and speed tests	average, percentiles, jitter + apply thresholds with % of deviations of installed base
capacity available and minimum required	per layer: sum of capacity + required + speed room derivatives
congestion	apply threshold for speed room + % of deviations & legacy definition

7.4.1. Cadence

The cadence at which KPIs should be collected is a balance between the speed for the business to react and the quality and work to collect these. In many cases a weekly collection provides this balance. Monthly and quarterly reports are then relatively easy to compile. In some cases, such as fault and failure monitoring, near real time is necessary.

7.4.2. Categorization

As discussed in previous sections, capturing metrics in a model that follows a universal, simplified architecture of the network has multiple benefits (see: fault management). It seems logical to capture KPIs in the same way. The example in Appendix A seems a good candidate to build such a framework.

7.5. A Framework

7.5.1. Data Models

These data models are templates for useful information in numerical form. There are a few basic types: logic(T/F), state (discrete), count (discrete numeric), measurement (continuous numeric), and some more complex measurements (such as complex I/Q data). Here are examples.

- Component state: up, down, more – translates to availability and reliability state

- Component capacity: assigned/allocated capacity (state change)
- Consumed capacity (real time): translates to capacity management for planning and engineering. This category includes functions including resiliency mechanisms – for example how many bit errors can be tolerated before the packet is lost.
- Component headroom: some network elements have limitations on their direct capabilities, such as transmit power. Treat these like spare capacity.
- Process success: can be T/F, but best to have measures of performance
- Deviation from target(s): includes quality – for example link quality at install may involve a measure of power level, with anything over a target is good.
- Defect counts: Like BER for example, counts of defects which translate into impact on quality or user experience when no resiliency is present.

7.5.2. Layering

Telemetry is raw data. Translation turns it into useful, interpretable information. This information needs to be transformed via a model into performance measures, and orientation information. For example, component states translate into availability estimates, and consumed capacity is compared to available capacity to determine a utilization estimate.

Performance measures need to be combined via a model into effectiveness measures. Additionally, orientation information helps to make decisions, based in part on performance measures and the effectiveness measure.

KPIs can come from effectiveness measures, or performance measures, or translations of these for specific purposes. They help determine the overall performance. Useful KPI have additional qualities that make sure they align to and measure performance of important things in effective ways.

The SCTE work group 'NOS WG8, network and service reliability', published a document that guides operators on translating network data and statistics into measures of performance and effectiveness that align to network and service reliability, the latter being complex. [5]

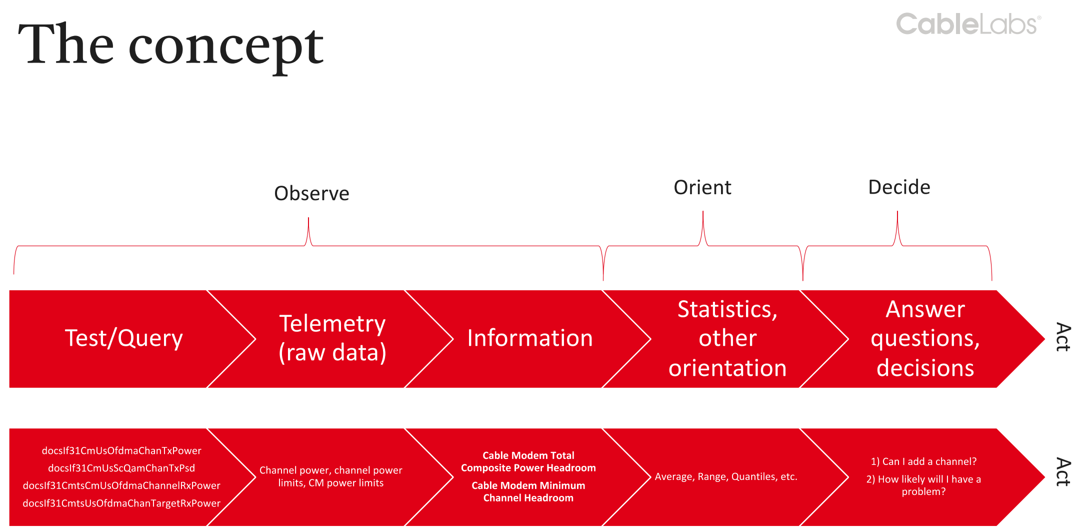


Figure 11 – concept of data layering

From a service perspective, communications serve applications, and those applications require throughput or goodput, appropriate latency, low enough jitter, infrequent packet loss, and communications that are almost always available. Throughput or goodput relate directly to bandwidth. Latency matters depending on the application, but network latency can sometime be managed by the application layer. Jitter is latency variation and can at times be managed by buffers at the application but can at times latency and jitter can lead to packet loss. That packet loss, which can also happen from network impairments and other issues, can lead to retransmissions or loss of data to the application, and that impacts the service experience. Availability of the communication, and the reliability of that communication for the session duration, impacts the service experience too. As a result, the service experience can be determined by a measure of effectiveness that is formed from performance measures that link these five requirements on the communication.

8. Conclusion & Recommendations

This paper discovers KPIs relevant to customer experience, performance and cost of operating networks through the analysis of business processes, The result is a relatively short list of recommended KPIs which largely resembles established HFC metrics.

KPIs rely on the availability of telemetry which emphasizes the relationship with the ongoing work to define this in the OOM workgroup. The use cases driving this development can be easily linked to proposed KPIs.

Besides telemetry, data from other sources is needed to complete all KPIs. In some cases, however, clever use of telemetry can reduce the necessity of external data by providing approximation. The proposed structure supports a natural evolution of telemetry and KPIs to grow as operations mature.

Standardization of underlying telemetry and data aggregation has many benefits such as efficiency in development and re-use of existing tools. It also enables benchmarking and supports cases where parts of the network or services are managed by different parties.

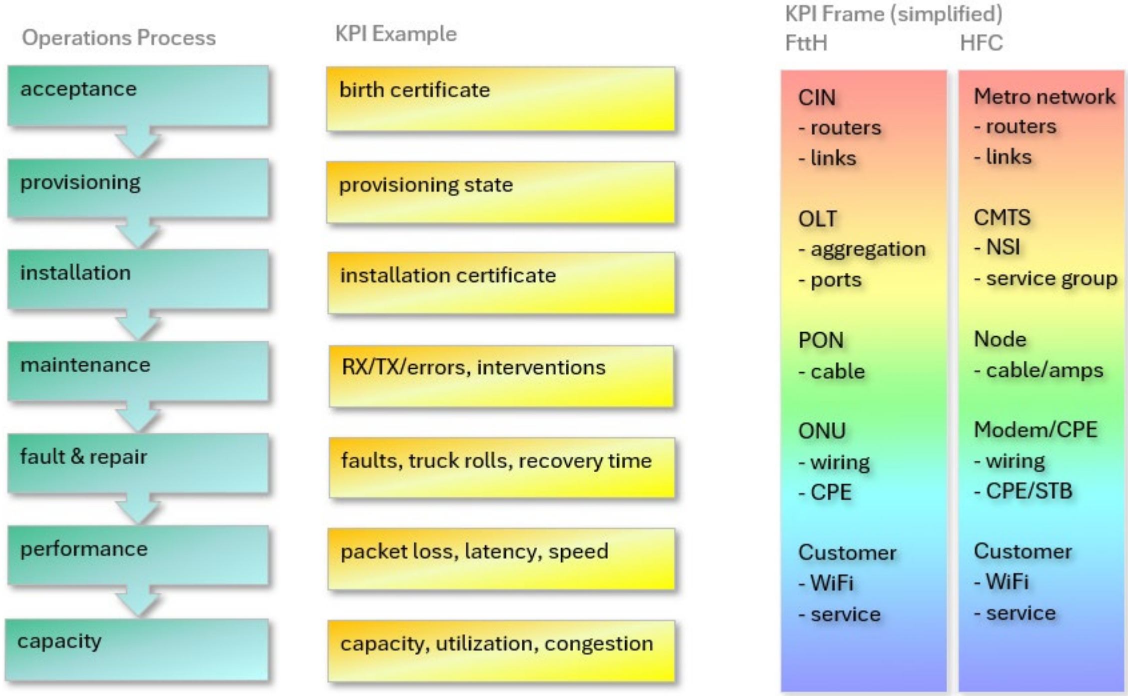


Figure 12 – KPI example and capturing frame

It is recommended to align KPIs and telemetry as much as possible across the industry and further collaborate to let the full model evolve with the business. This paper is intended to contribute and inspire to the development.

Appendix A – Fault Categorization

An example of fault categorization.

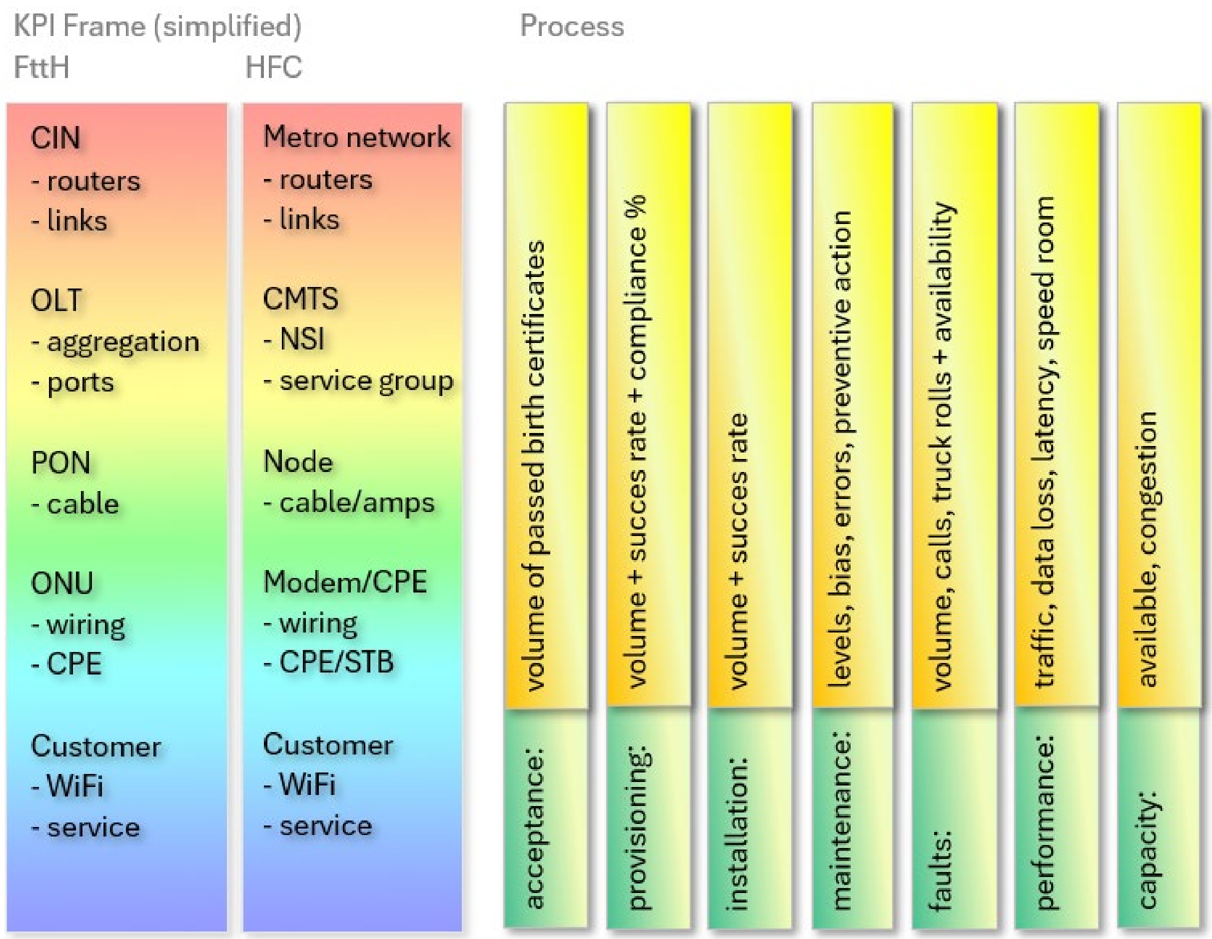
Faults / truck roll fix codes are counted vertically by network layer. On the horizontal axis, these can be divided into network sections and if required sub-sections, for example related to the source (telemetry, NOC, call center).

In the resulting matrix, different crosspoints provide KPIs to manage various processes and departments. It supports cases where different layers are managed by different owners (e.g. wholesale/wholebuy or ServCo/NetCo).

Truck roll resolution categories	HFC	FTTH
Network		
Transport (D1) (incl. fiber node/OLT)	Master Telecom Center CMTS Distribution HUB Fiber Optic Transport Fiber Node	CIN OLT Point of Termination (PoT) Fiber Optic Transport Point of Presence (PoP)
Distribution (D2) (incl. tap/POC)	Coax from Node Amplifiers Tap	Fibre from PoP Splitters Point of Connection (PoC) Building Entry Point (BEP)
Tap to Home (D3) (incl. 1st wall plate/ONU)	Coax Drops Tap to Home	Fibre (BEP-)POC to home
in-home		
Home network wiring	Home Coaxial Wiring	Home Fiber Wiring
CPE Hard-/Software	CPE Hardware-Software	CPE Hardware-Software
Home data network /WiFi (network after CPE)	Home Data Network /WiFi	Home Data Network /WiFi
avoidable		
Customer Education	Customer Related Education	Customer Related Education
No Fault Found	No Fault Found	No Fault Found
Not Home Cancel at Door, Reschedule	Not home	Not home

Appendix B – Simple KPI chart

A simplified chart to summarize the various processes and examples of KPIs towards the KPI frame.



Abbreviations

HFC	Hybrid Fiber & Coax
FTTH	Fiber To The Home
KPI	Key Performance Indicator
OOM	Optical Operations and Maintenance
EMS	Element Management System
SMART	Specific, Measurable, Attainable, Relevant, and Time-Bound
dB	decibel
OLT	Optical Line Termination
PON	Passive Optical Network
ONU	Optical Network Unit
OTDR	Optical Time Domain Reflectometer
TX	Transmit
RX	Receive
ID	Identifier
CPE	Customer Premise Equipment
CIN	Converged Interconnect Network
X-GEM	XGS or GPON Encapsulation Method
DOCSIS	Data Over Cable Service Interface Specification
AI	Artificial Intelligence
NOC	Network Operations Center
bps	Bits Per Second
Mbps	Megabits Per Second
FEC	Forward Error Correction
X-GEM	XGS-PON / G-PON Encapsulation Method
HEC	Header Error Check
PNM	Preventive Network Maintenance
MER	Modulation Error Ratio
CMTS	Cable Modem Termination System
T/F	True or False
I/Q	in-phase (I) and quadrature (Q) components of a signal
SCTE	Society of Cable Telecommunications Engineers
SLA	Service Level Agreement
CIR	Committed Information Rate
ServCo	Service Providing Company (separated from Network Company)
NetCo	Network Providing Company (providing access for ServCo)

Bibliography & References

1) Telemetry definition: Wikipedia
2) Telecommand definition: Wikipedia
3) The speed triangle: SCTE EXPO 2022
4) https://www.clearpointstrategy.com/blog/18-key-performance-indicators
5) SCTE 290, "Service and Network Reliability Measurements and Use Cases," 2024