www.pipelinepub.com Volume 16, Issue 7

# Tackling the Burden of Poor Data Quality with Tagging

By: Jan Ulc, CROSS Network Intelligence

The 5G era brings transformation towards a new, service-oriented architecture for operators globally. The new network architecture is designed to enable the support of an increasingly wide range of services, with an equally wide range of different performance requirements. It's also supposed to introduce new dynamic capabilities, so that services can be composed, delivered and orchestrated on the fly, before being shut down or recycled. And, lest we forget, 5G isn't just for mobile: this new architecture will become the foundation of fixed as well as cellular networks.

It's a big vision for operators. To achieve this level of agility, they will have to transform their operations and processes. Simply acquiring the ability to spin up new services dynamically won't be enough, operators will have to be able to ensure that they deliver them, at the right time and place, and with the right experience required.

This presents challenges because customer experience is directly linked to service quality and performance. In turn, this requires careful orchestration of the resources on which services depend. Operators must know what resources they have available, where they are and how they are being used at any given moment. These resources include services, transport infrastructure, physical assets, information regarding their location and routes such as wayleaves, and metadata such as street addresses, as well as virtual elements that may be required when a service is composed

And, they not only need to know this information, but they also need to be able to obtain it, when required and in real-time, so it can be shared with appropriate business processes. That's because a key goal for the next-generation network is to enable automation. In this context, automation means that a service could be ordered by a customer, with the necessary actions taken to deliver and manage it being enabled automatically with zero-touch intervention.

It's not just customer orders, however. A series of conditions or events could also trigger the establishment of a new service or modification and extension to a live session, dynamically. Again, the orchestration and service platform must be able to check what elements and resources are required to fulfill these demands, whether they are available, and then to instantiate the change by composing the necessary elements in the right way. Operators need a dynamically updated, real-time view of resources, available as a complete picture: a single source of truth for their networks.

To understand and to obtain this complete picture requires accurate, high-quality data that is accessible to all business systems, as well as a consolidated record of all such assets. This must all be consolidated into a single data model—in other words, a comprehensive network inventory system.

# Data problems

The problem with this is that few operators have access to such data. Worse, few also have access to the single, consolidated inventory of their network assets that is required. Instead, many have multiple, disconnected inventory systems, populated with incomplete and poor-quality data. As a result, plans to deliver on the 5G promise and dynamic service orchestration are likely to be undermined.

The first step is to recognize that the problem exists. It turns out that failure to recognize this problem is surprisingly, if unintentionally, common. Many operators have absorbed legacy practices and systems, making do with different silos of data, even as they grow their network and subscriber base. This situation often results from mergers and acquisitions. Because until recently many operators had manual processes for activating and delivering services, the problems of such approaches were not apparent.

And, even if the difficulties that such legacy practices have caused are known and understood, the task of rectifying them can often seem too huge to tackle in the face of other network evolution and transformation initiatives. This is the well-known problem of inertia: we understand the problem but cannot take action, because we fear the consequences or think that it will demand too much effort.

Finally, data may well be available, but it may be scattered across different silos, in different databases and may also be incomplete and imperfect. Some operators have paper-based records or an array of spreadsheets, so the data available exists only in different, incompatible formats. Making sense of this can be time-consuming.

The data needs to be brought together, but the imperfections also create a significant obstacle. What’s needed is a means of processing this data, regardless of any imperfections, so that operators can embark on the task of transformation and inventory data consolidation but without the need for a seismic approach.

This looks strikingly like a problem of “irreducible complexity.” Only, just as the answer to the question “what use is half an eye” is “a lot more than no eye at all,” so there is an answer to this conundrum. Imperfect data is still data, regardless of its perceived quality.

# Data quality, metrics, and tagging

In our opinion, there aren’t hard and fast rules regarding data quality. Assessment of the utility and completeness of data can be made on a more or less subjective basis, subject to some basic principles. The trick is to come up with some useful metrics that provide a benchmark and baseline for action, provided that a meaningful scale is created. Simply by creating a range—with 1 as the highest quality and 5 as the least—tagging will enable users to assign scores to data, according to their own criteria. A few examples will serve to illustrate the point.

Let’s consider a fiber that needs to terminate on a port in a specific cabinet or device. This is required for the delivery of a fiber service to a particular customer (a data point on its own) at a particular location (another data point). If we know where the cabinet or device is, we can give that data point a high score—say 1. If we know the fiber, but we don’t know on which cabinet or device it is terminated, we could give this data point a quality score of 5. If we do know which cabinet or device, we could assign a score of 1. Similarly, if we don’t know which port on the device, we may use a score of 3, and so on. This not only allows us to capture the data but also to generate a map of its quality, allowing it to be used immediately.

The same logic can be applied to a cell site and fiber rollout program. We may know where the cell has to go (1) to give optimum coverage, but we may not know where the nearest fiber duct is (5), which is essential for delivering the backhaul capacity required for 5G performance. We can use these to categorize tasks, understand the difficulty in completing them and complete the overall picture of network resources. Once we know where the fiber duct is, we can upgrade the quality score to 1.

In this way, even though we are using arbitrary values, we are assigning them logically. Completing this exercise gradually allows a baseline to be created. It allows all resources with a high-quality value (1) to be used immediately while allowing those with poorer quality values to be identified. This latter point is crucial because it means that a program to drive continuous quality enhancements can be initiated and followed, with clear targets in mind. However, it can be pursued when time and resources allow.

Such a program can be prioritized, too. After the initial exercise, further rounds can be launched. It could be addressed, for example, to all data points with a value of 5, then to all with a value of 4 and so on, so that, iteratively, overall data quality can be enhanced, step by step from worst to best, or according to any preferred schedule.

By tagging data, it can be rendered immediately useful, even if caveats apply. Imperfect data can then be imported into a new, consolidated inventory system and made available to other processes, from the outset. This also lowers the barrier for new projects. If data quality and disaggregation have been a blocking point, resolving the issue with a clear, quality tagging process means that project delay can be eliminated and costs reduced.

# Enabling an agile approach

It means that operators can adopt a more agile approach even during transformation projects. Because quality can be enhanced iteratively, teams can migrate towards a new, consolidated data model with a single inventory of network resources, working on the principle of “good enough is good enough,” so that they can immediately reap benefits. Of course, further benefits will be unlocked through time, but even half an eye lets a little light in. The resulting agility will underpin future efforts to capitalize on new-generation networks and services.

The questions of data availability, accessibility and quality are of fundamental importance. In order to deliver the performance and agility that 5G and the new network architecture promises, operators must be able to support dynamic service creation, orchestration, and delivery. They need to deliver to both customers and to support new business partnerships. In addition, operators are shaping up to support a new range of connected devices, which may be delivered directly or by third parties. In either case, they will become new resources to be considered within inventory management.

The foundation of this approach is network inventory: a consolidated record of all network resources and assets that is accessible to other solutions. The bricks and mortar of the inventory are the data that is imported, often from disparate sources and with variable quality. Many operators are challenged by imperfect and poor-quality data. But, by adding data quality metrics, the burden of poor-quality data can be eliminated. Migrating to a new, consolidated network inventory—an indispensable tool for future operations—need not be a challenge. The innovation of data quality tagging solves this problem, delivering clean data and providing the foundation of agile business and operational processes to support 5G and a new service model.