Executive Briefing

MONETISING 5G: HOW IT/OT CONVERGENCE WILL ENABLE NEW BUSINESS MODELS

Real-time data processing is a key pillar for operators looking to achieve revenue growth with 5G-enabled services. They must implement it internally, and externally with their enterprise clients. This paper explores four key value-unlocking use cases where IT and OT converge.

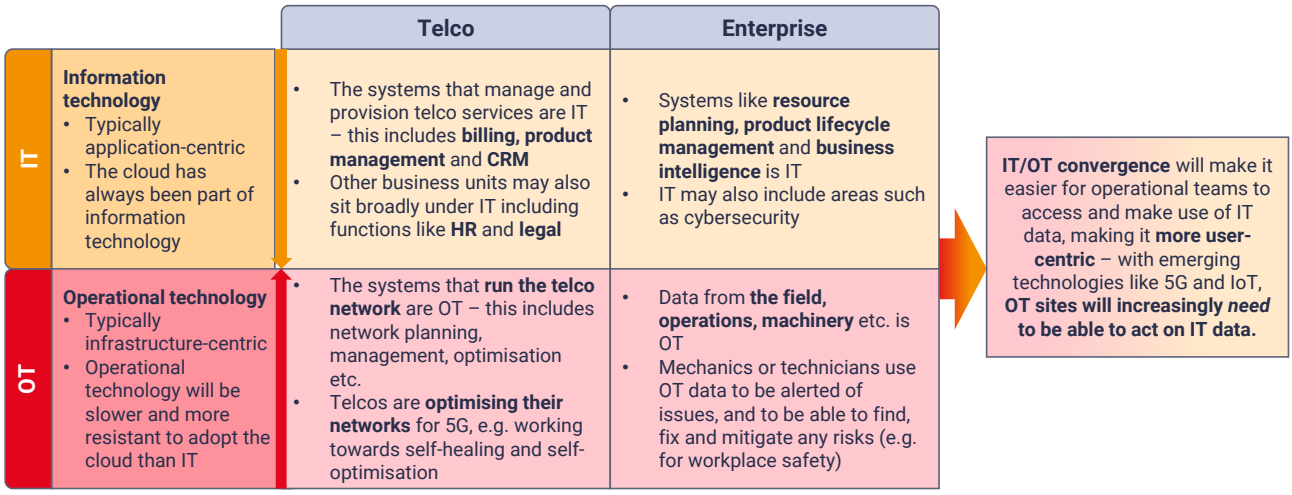
Tilly Gilbert, Senior Consultant | tilly.gilbert@stlpartners.com | March 2021

Reah Jamnadass, Consultant | reah.jamnadass@stlpartners.com |

Executive Summary

- Real-time decision intelligence is critical to telecoms operators' future, both in increasingly internal efficiencies and enabling them to deliver new revenue-generating solutions to their customers.
- The nature of changing customer requirements, particularly in light of 5G massive machine-type communications (mMTC), means OT and IT data must increasingly be handled together.

Figure 1: IT and OT convergence is a trend impacting both telecoms operators and their enterprise customers



- In this paper, we identify and explore four use cases that need to evolve to unlock decision intelligence in real-time and handle the convergence of IT and OT data:
  - Policy management: must evolve to handle 5G network slicing
  - Rating and charging: must evolve to handle M2M communications
  - Fraud prevention: must evolve to support IoT use cases
  - Customer monetise customer engagement points: must evolve to better manage investments such as 5G rollout
- To capitalise on these use case, telecoms operators and others within the ecosystem should consider these practical steps:

1. Do not wait for 5G to explore the value they can unlock with low latency data processing. There are plenty of opportunities both in terms of internal efficiencies and revenue generation that can be achieved with existing network infrastructure and moving early will ensure a strong platform for 5G-enabled services is created.

2. Consider moving to cloud-native networks and IT – by doing this, IT and networks can interface more closely, and siloes can be broken down between the network and the monetisation engine supporting it. A cloud-native architecture can also better support microservices, critical to fostering agility and achieving faster time to market.

3. Ensure that solutions can support agile product development and innovation – since we are not seeing one obvious "killer 5G use case" it is likely that there are many, in fact, and operators need to be able to innovate faster and more efficiently than before to capitalise on these opportunities.

4. When aiming for low latency, do not compromise on reliability, resilience, and predictability of the system. For many use cases, both internal to the telco and external to their customer, reliable low latency is often more important than low latency itself.

5. Explore data platforms that can facilitate low latency decision intelligence. The amount of data is only going to increase as 5G rolls out, so operators will want to make decisions that will future-proof them.

Table of Contents

Table of Figures

Introduction: the unspoken reality of 5G monetisation

The "5G opportunity" is much hyped by almost everyone within the telecoms industry. Its rollout promises to unlock huge potential revenues for telecoms operators and their technology partners, particularly within the enterprise domain. However, for this to be possible, operators must make changes and invest in new technology. Telcos across many markets are already investing in their networks, by acquiring spectrum and improving the underlying hardware and software to be ready to deliver 5G. In other words, they are investing in their "OT"; operational technology – the systems that run their core business, the network.

However, it is not just OT where operators will need to make changes. There is a pressing need for operators to evolve and update supporting IT infrastructure, if they are to stand any chance of converting their network investment into a revenue opportunity and competing in the evolving telecoms landscape. The reality today is that many operators' BSS systems rely upon 10+ year old architecture which cannot deliver the highly automated, highly flexible system that is now required.

Without considering their IT systems alongside their networks, and ensuring that both are fit for a 5G world, operators will struggle to support:

- Improved customer experience, such as more personalised product packages and self-service portals;
- New charging models, such as those that will support mMTC numbers of IoT devices;
- New services, such as 5G network slicing;
- New products and product delivery models, such as bundling network services with solutions delivered by an ecosystem of partners;
- New innovation cycles, where product development take days/weeks rather than years.

The challenge comes not just from having to improve network and IT systems independently, but in ensuring that they efficiently and consistently interface with each other. Previously, a mediation layer has enabled information such as call data to be imported into a billing system. Going forward, these data types will increasingly need to be aggregated and processed together, in real-time. It is only through doing this for themselves, with data from their IT and network domains, that operators will gain the skills and technology they need and prove their pedigree to enterprises in other industries undergoing similar journeys. This will enable operators to be perceived as valuable partners and enablers for industrial IoT solutions converging IT and OT data within the manufacturing industry, for example.

In this paper, we explore what needs to change to enable this to happen, leveraging insights from conversations with industry experts including telecoms operators and their IT partners.

Real-time decision intelligence: the future for telco operations

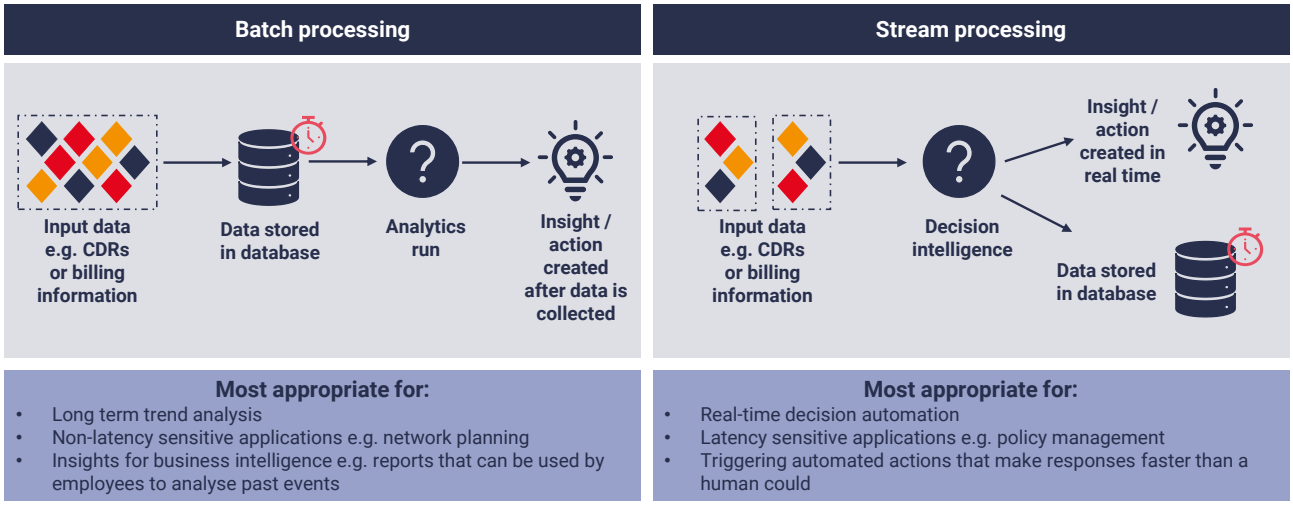
As operators seek to become highly efficient, highly automated companies, the importance of data collection and processing happening in real-time is only increasing. As they begin to commit to 5G rollout, the demand for immediate data-driven insights and decisions is growing:

- Within networks/OT: to handle the complexity of 5G networks, operators are increasingly focused on closed-loop automation to create a network that can be self-optimised, self-provisioning, self-healing, etc. To achieve this, actions will need to be triggered in real-time, responding to network demand, customer SLAs and so on.

- Within IT: customers demand services at their fingertips, e.g. they now expect that changes to their bill will be made and shown to them in real-time. In tandem, operators will offer increasingly bespoke SLAs to enterprise customers, e.g. to offer application-specific mission-critical M2M services. This will require complex and reactive policy engines that work to guarantee that SLAs are being met and the required quality of service (QoS) is being delivered. These too will require real-time data processing and decision-making.

To evolve from post event analytics to real-time processing and decisioning, operators will need to move away from batch processing of data, and instead use stream processing. Stream processing means that decisions are made as input data is ingested, rather than waiting for a batch of data to be stored before a query is run. Within IT, some processes have already moved to stream processing, e.g. for charging and rating, but others continue to be done via batch processing. Within the networks, stream processing is the norm, but accessing real-time data from legacy network systems can be a challenge.

Figure 2: Stream processing is critical to being able to unlock the value of real-time insights

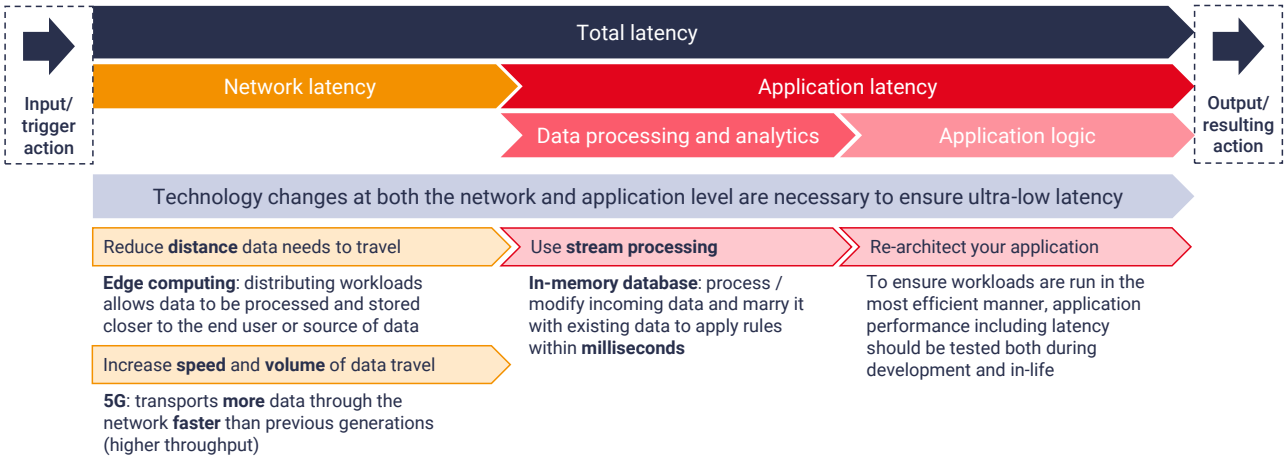


Source: STL Partners

Stream processing is different to, but complimentary, to other key data-driven trends such as the use of analytics on big data. While big data analysis draws insights from many data sets often collected over a significant period of time, and identifies trends and patterns within that data, stream processing is all about immediate and correct intelligent decisions which in turn evoke appropriate actions.

The trend towards stream processing is not only about low latency decision-making; it can also help telcos to manage the vast quantities of data that they are handling. Since data collection has become a 24/7/365 phenomenon, it makes sense for processing to happen as the data comes in, rather than after the fact. As the CTO of an advanced analytics platform put it, stream processing can reduce the cost of storing data and can ensure it does not build up into something slow and unwieldly. He likened it to the dangers of a traffic jam: *"when there is a crash on the motorway, you always try and keep one lane open to make sure traffic doesn't build up".*

Figure 3: A combination of factors impact latency, including both network and application processes



Source: STL Partners

Real-time, automated actions rely upon short end-to-end latency. As Figure 3 indicates, operators can look to reduce the response time in multiple ways within that window. One major trend to achieve this will be the move to more distributed architectures, to physically run decision intelligence on a server close to where the trigger event data is generated. As both IT and OT deploy stream processing, a distributed architecture will enable both of these domains to exist on servers close to each other (or even on the same physical infrastructure) – this will reduce or even remove the impact of network latency entirely, enabling short latency from input to resulting action. An example of this would be the rating and charging for a 5G slice being done at the enterprise site itself.

In the next section we will dive into the importance of real-time decision intelligence and the challenge of IT and OT convergence within four specific telco domains. We will also discuss how these domains will help operators to realise their ambition of 5G monetisation.

Proof-points: bridging the IT and OT divide through real-time decision intelligence

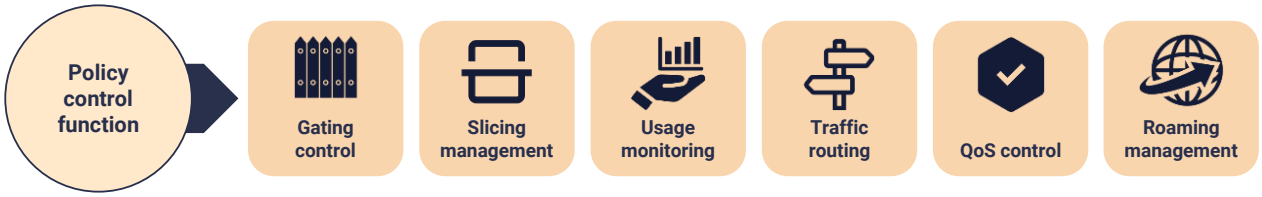
Policy management

Real-time policy control will be a key enabler for operators seeking to both monetise 5G and handle its increasing network complexities. Policy management can encompass account level rules, QoS (quality of service) policies and network optimisation rules as well. Essentially, the policy management function determines how the network will behave to serve each customer's need. Acting on a set of rules that considers factors such as customer behaviour and usage, the function allows operators to control how their network resources are configured to best deliver services. But with 5G, policy management functions will need to adapt to both meet and exploit changes in network usage:

- As the network becomes more complex and more devices (with variable data requirements) are connected, policy engines will need to control vast amounts of data traffic while ensuring a high QoS to meet customer needs;
- As operators seek to secure a return on their investments with 5G, policy control will require more advanced processing and decision-making capabilities to control network behaviour according to customer usage at a more granular level, such as providing a different QoS to different customers or even network slices;

With stream processing, these decisions on network behaviour can be based on real-time processing of incoming data, in addition to pre-determined insights, enabling more dynamic control of network resourcing.

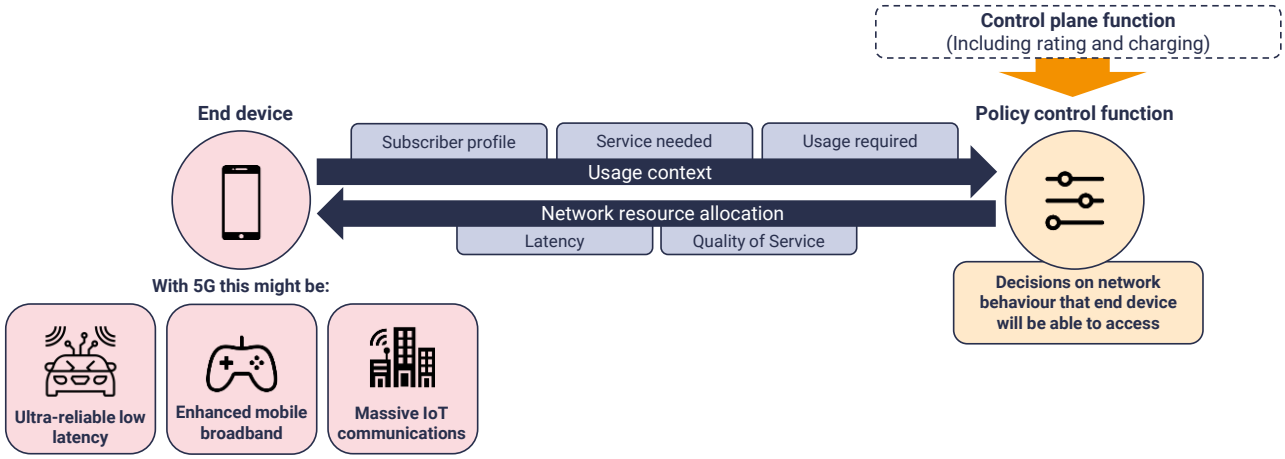
Figure 4: Policy management is responsible for several functions that enable operators to monitor and control the behaviour of their networks



Source: STL Partners

In this (5G) context, real-time data-processing for policy management will not only enable operators to manage their network resources more efficiently and cost-effectively but can also provide granular visibility over *how* their networks are being used to enable intelligent decision making rapidly. This will make it easier for operators to manage and allocate resources as they seek to monetise 5G and enable use cases with stringent network requirements (e.g. low latency). To do so, policy engines must be able to ensure a consistently high QoS for end customers. This is something that will become more critical and difficult to do, due to increased network traffic with 5G, which has more than double the number of QoS indicators compared to 4G.

Figure 5: Real-time policy control will make it easier for operators to manage increasingly complex (5G) network functions and resource allocation



Source: STL Partners

For example, as ultra-reliable low latency use cases become more commonplace (e.g. in industries such as healthcare and manufacturing), it will be critical for policy engines to ensure that SLAs are *always* met (e.g. ensuring there is no buffer-time for a video stream supporting a surgical procedure). The ability to offer these 5G-based customised SLAs could be a key enabler for operators seeking to monetise their 5G network. To achieve this, operators will need to ensure the end-to-end management of their network is highly automated across all domains: network optimisation, re-routing and provisioning.

Beyond 5G applications, network slicing will have a particularly big impact on the complexity of policy management. Operators will need to manage QoS and resource allocation on a per-slice (and potentially even per-session) basis. Again, this will depend on both real-time and granular processing of data to ensure the network can intelligently distribute resources across slices to manage individual and diverse slices.

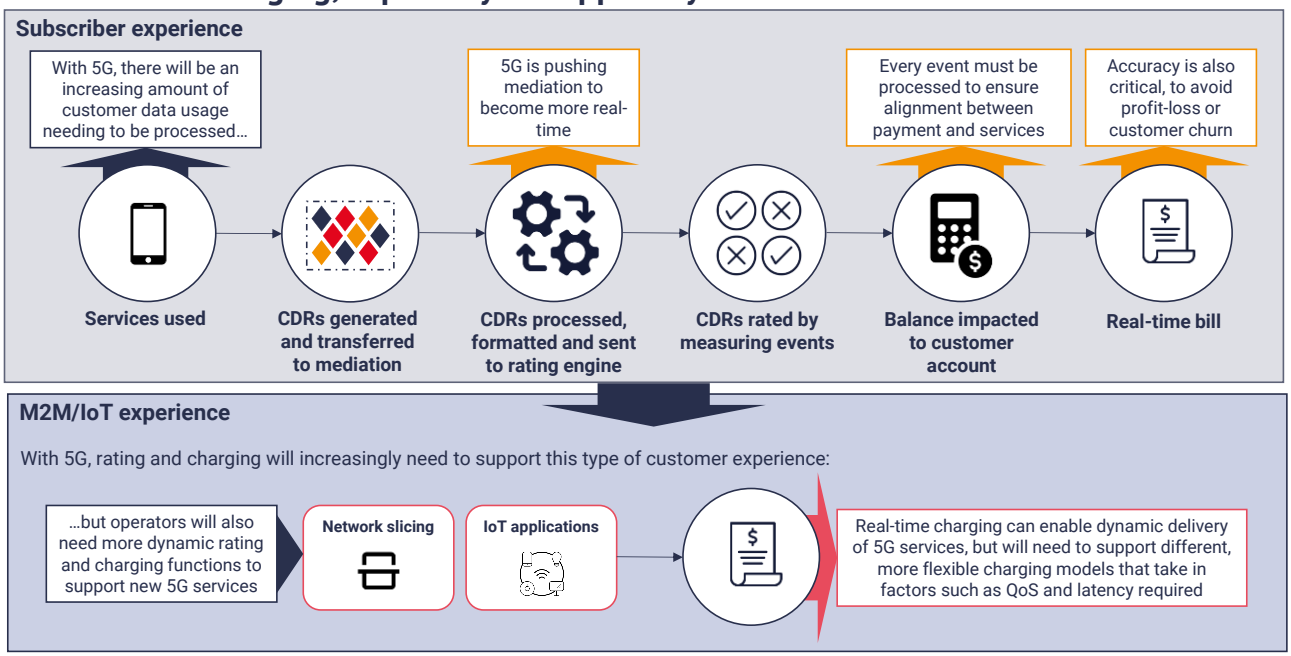
Rating and charging

Rating and charging have some of the most stringent latency requirements within the IT domain. At a basic level, they ensure that customers are receiving the exact services that they have paid for: operators do not want to withhold services that customers are entitled to receive, but they also want to avoid customers accessing services they have not paid for. This system, involving the real-time stream processing of OT data will be increasingly important with 5G as:

- Customer data usage will increase, especially with dynamic IoT environments that will produce much higher volumes of network traffic
- There will be new business models, requiring dynamic and more flexible rating and charging models: as operators seek to enable new products and services with 5G (e.g. IoT), their charging systems will need to be able to configure accordingly

For example, managing use cases such as 5G network slicing brings additional charging complexities – devices may connect to different slices according to latency needs (e.g. to a uRLLC slice when needed) and charging models will need the sophistication to accurately reflect this; as one APAC operator commented: *"Can BSS systems handle the granularity and complexity of 5G use cases like slicing? How do you charge? Per slice?"*

Figure 6: Ultra-low latency processing is becoming increasingly critical for rating and charging, especially to support dynamic 5G and IoT services



Source: STL Partners

Equally important to this use case is the *reliability* of real-time data processing. 10-20ms is the optimal latency for rating and charging data to be processed, regardless of how many transactions are taking place at the time, or whether there is network congestion. Average latency in the field is a key metric, in addition to the best-case latency that can be achieved in the lab. With 5G, there is only going to be more data that needs to be *reliably* processed, as well as quickly.

To ensure this reliability, data processing technology needs to be highly accurate for rating and charging. With big data, if the aim is to find patterns, some data loss is not a big issue – 99.5% accuracy is most likely sufficient for reliable conclusions to be garnered based on the data. However, this is not the case with rating and charging – any loss of data or inaccuracy can result in the wrong amount being charged to the customer, which could result in a revenue loss (undercharging) or could impact customer satisfaction and lead to churn.

However, real-time and reliable processing of data for charging will not only put operators in a good position to handle changes to network usage with 5G but could also allow them to offer more dynamic charging models to customers. For example, especially as IT and OT data continue to converge, operators may be able to exploit more granular intelligence of contextual data usage reflecting *how* the network is being used, tying in with the policy management function, and enabling more dynamic

charging models. For example, operators might be able to charge according to multiple factors such as QoS required (e.g. mission-critical applications depend on ultra-reliable as well as ultra-low latency) and complexity of environment (e.g. whether multiple slices are being used).

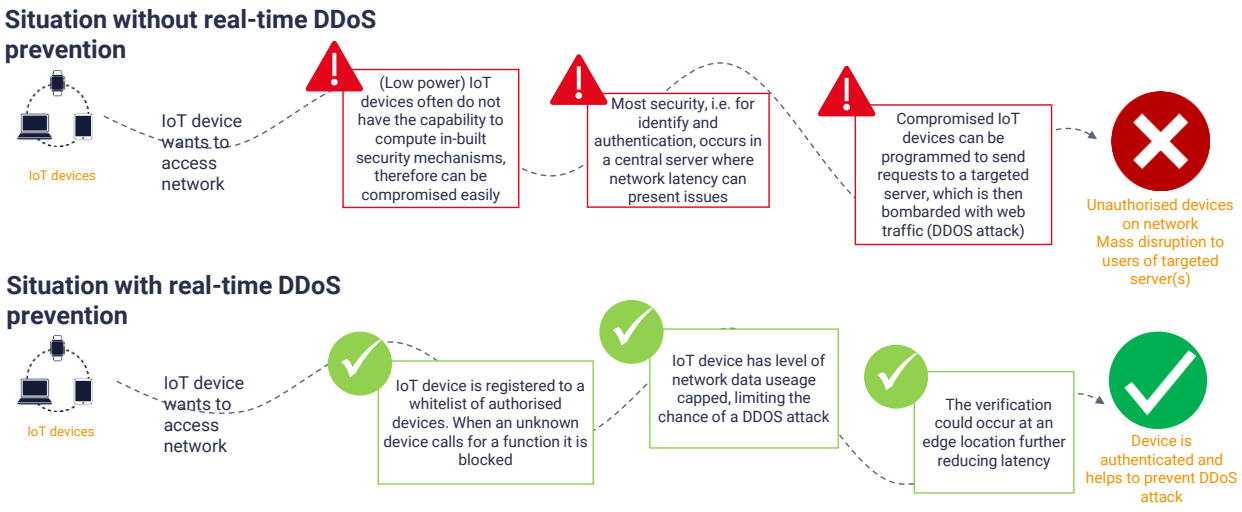
Fraud prevention

The impact and attack surfaces for fraud are growing rapidly. Nowadays, operators not only need to consider subscriber fraud, but also the impact of the rapid increase in the number of connected devices under the operator's domain. These M2M communications will bring the domain of fraud prevention from a purely IT use case into the world of operational data.

There is a window of less than 200ms in which a telecoms operator needs to capture an incoming event, use a set of rules or ML algorithms to determine whether it is a fraudulent event, and then act to block the fraudulent activity from taking place. Achieving this means operators can move from identifying and managing fraudulent activity after the fact, which has a negative impact on subscriber experience and can cause revenue leakage for the operator, to being able to prevent it from happening in real-time. Handling the sheer volume of processes in which fraud could occur means automation is critical.

Automation is all the more important when there are no humans as part of the transaction that may be able to identify or intervene if they recognise unusual behaviours or activity. In the case of IoT devices which may be sharing or transacting valuable customer data, being able to detect these anomalies automatically and quickly is highly important. This requires stream processing. One example of this highlighted by the CTO of a stream processing platform is being able to block IoT devices that should be stationary (such as a connected vending machine) if their IP address and location data indicates that they have been moved. Another example would be the importance of handling DDoS attacks – the ability to block the compromised device in real-time here could be the difference between there being no noticeable impact to major outages of websites and services.

Figure 7: Real-time decision intelligence will enable operators to offer secure and resilient IoT solutions



Source: STL Partners

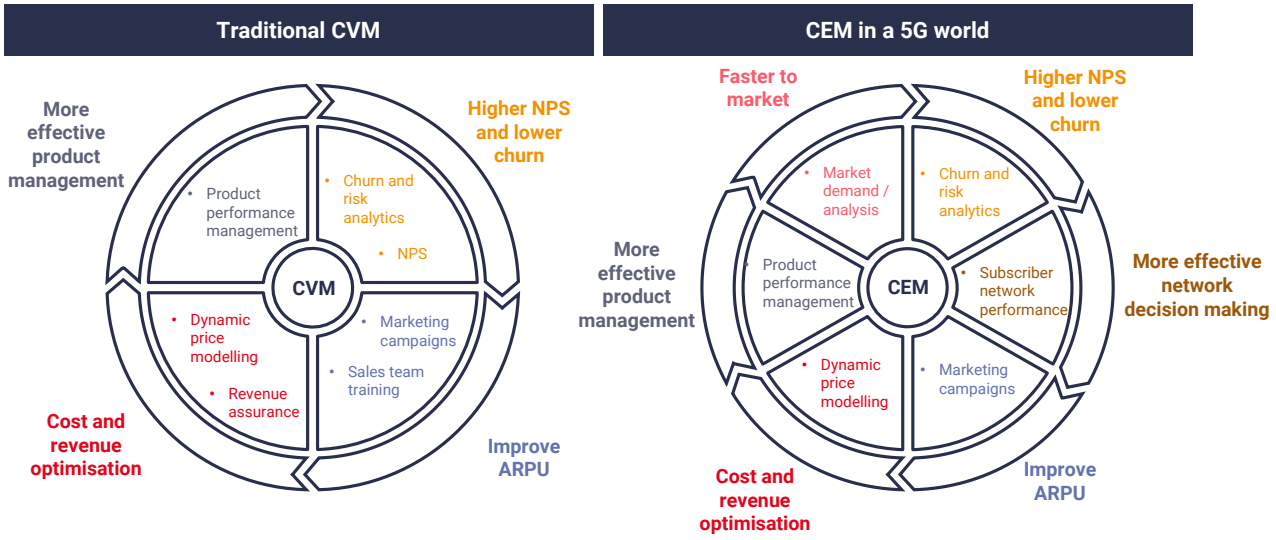
Customer management

Particularly in the consumer space, where 5G plans are starting to offer unlimited data bundles, the world of customer management (encompassing customer value management, customer experience management and customer lifecycle management) is changing. For many, churn remains relatively low and instead the focus is on increasing ARPUs (average revenue per user) while also better managing both CAPEX investments (e.g. in new networks) and OPEX investments (e.g. in sales and marketing programmes).

To do this, operators are looking to understand a more holistic view of the customer satisfaction – combining account IT and BSS data with key network performance data too – to determine the service the customer experiences and pre-empt issues. Input data sets could include: network data (PCRF, Voice, MME), IT data (tariffs, billing, website data), customer care data (number of complaints), customer profile, customer NPS, churn predictions, etc. Doing this can help to guide operators to prioritise network investments that have the biggest impact on customer satisfaction – this will be particularly important as operators look to roll out costly 5G networks.

For a Middle Eastern operator, this transition to holistic customer management has indicated a change in the organisation's thinking. Firstly, it has meant that the exchange of data between networks and IT has become a two-way process; network engineers both provide data that impacts customer experience scores, and use customer management data to inform their networking decisions. Secondly, it has accelerated acceptance of customer management as a horizontal use case that must span across all the operator's business units, supporting services from roaming to value added services and beyond.

Figure 8: Customer management will become a use case that both ingests and outputs data for use within IT and OT domains



Source: STL Partners

This evolution includes changing how insights are generated through decision intelligence, as well as aggregating new data sets from the network. Enabling real-time communication with customers, such as an offer of a top-up when a customer is watching a film but about to run out of data, can impact key metrics such as cost of retention and churn. As operators offer 5G-enabled services, customer management will have to handle increasingly complex scenarios, particularly as operators look to partner and enter into revenue share agreements with those partners. As services are requested and provisioned in real-time, customers, enterprise and consumer alike, will expect telco systems to reflect these changes instantly and partners will expect these systems to give them real-time visibility and oversight too.

Conclusion: how to make this happen?

With the advent of 5G-enabled services, operators will increasingly need real-time solutions that can handle different types of data. First and foremost, this will enable them to break down internal siloes, improve customer experience and ensure they are able to deliver new use cases where M2M communication is critical. Achieving this will also demonstrate that operators have the skills and capabilities to do the same for their end-customers, including becoming a trusted handler of both IT and OT data from their enterprise customers.

Operators should consider the following practical steps in their journey to achieve this:

1. Do not wait for 5G to explore the value they can unlock with low latency data processing. There are plenty of opportunities both in terms of internal efficiencies and revenue generation that can be achieved with existing network infrastructure and moving early will ensure a strong platform for 5G-enabled services is created.

2. Consider moving to cloud-native networks and IT – by doing this, IT and networks can interface more closely, and siloes can be broken down between the network and the monetisation engine supporting it. A cloud-native architecture can also better support microservices, critical to fostering agility and achieving faster time to markets.

3. Ensure that solutions can support agile product development and innovation – since we are not seeing one obvious "killer 5G use case" it is likely that there are many, in fact, and operators need to be able to innovate faster and more efficiently than before to capitalise on these opportunities.

4. When aiming for low latency, do not compromise on reliability, resilience and predictability of the system. For many use cases both internal to the telco and external to their customer, reliable low latency is often more important than low latency itself.

5. Explore data platforms that can facilitate low latency decision intelligence. The amount of data is only going to increase as 5G rolls out, so operators will want to make decisions that will future-proof them.

STL PARTNERS

Research

Consulting

Events