cpl.thalesgroup.com

THALES

A New Trust Model For The 5G Era

White Paper

Contents

5G promises to be rich in new business models for users and the ecosystem key players (telecom operators, network & cloud vendors, system integrators...). Some of the requirements that exciting new 5G use cases impose on the storage, compute and network domains introduce sizable new risks to the confidentiality, integrity and availability of enterprise data. This White Paper provides Thales' recommendations to the 5G ecosystem on how these new challenges can be navigated to build a new trust model for the 5G era.

The 5G Opportunity – The View From the Enterprise

5G promises to put a phenomenal set of capabilities into the hands of public and private enterprises across industry verticals. Compared with 4G's one-size-fits-all approach, the 5G ecosystem will give organizations unprecedented opportunity, flexibility and choice in the networking tools they can use to capture and store data, and then glean insights from it to drive their digital operations.

The 5G access portfolio will enable IoT devices to generate and exchange a variety of different types of both high and low value data at scale and at low-cost points that 4G alone can't reach. Every generation of cellular brings a higher capacity radio access network (RAN) that allows the same data to be transmitted faster. 3GPP's 5G New Radio (NR) does that too, but this time 5G also leverages additional technologies like Wi-Fi, and the 5G massive IoT segment builds upon the new NB-IoT standards whose performance characteristics are optimized for a variety of new IoT use cases.

As telecom operators embrace cloud-based – even cloud hosted – operating models and edge computing, 5G will provide an unprecedented platform for automated, zero-touch operation of the network, customization of network services and superior performance of enterprise applications. The network slicing capabilities of 5G will enable enterprises or service providers to have their own autonomous 5G networks, each one customized to their unique requirements and backed up by an SLA.

Another remarkable aspect of 5G will be the ability to drive ultra-low latency down to a few milliseconds which will require a highly distributed network architecture. There's nothing in 3GPP standards mandating that today's 4G network should be largely centralized and manually provisioned as most still are. It's just that the key enablers of distribution – Network Function Virtualization (NFV) with an automated and orchestrated NFV Infrastructure (NFVI) and Multi-access Edge Computing (MEC) – are only now starting to roll out at scale. Automation, orchestration of a more distributed network remains optional for 4G but it's critical for 5G.

The greatest opportunity available to enterprises from 5G is the insights derived from combining analytics, machine learning, and Artificial Intelligence with the unique capabilities of the 5G network. Most businesses will want to leverage 5G to capture, move and store data. Sustainable competitive advantage awaits those businesses that then apply analytics to the data in optimal locations to drive automated events in day to day enterprise operations and in the daily lives of consumers.

This white paper touches on the opportunities of 5G. What it focuses on, though, is the fabric of trust that needs to underpin the 5G ecosystem. In particular it focuses on the supporting requirements in 5G system security and data protection; IoT services; authentication and authorization of end devices; and the connectivity and software monetization that will be needed to fully capture those opportunities and convert them from theory into practise.

A New Trust Model for 5G Networks

5 Points of the Thales 5G Vision

I. 5G system security and data protection

II. Device lifecycle management

III. Device authentication and authorization

IV. Connectivity and network slicing security

V. Software monetization and NFV

1. How to answer the new data protection challenges

While the 5G era is brimming with new opportunities, it is also loaded with new risks. In the case of all previous cellular generations, including 4G, the greatest risk to consumer and enterprise data is still the same: confidentiality being compromised by interception of voice or data communications at the transport layer or of unauthorized access to a universally available set of essentially one-size-fits-all network services.

Leveraging end device authentication as well as encryption across the RAN, transport and core, 4G has extended the cellular industry's excellent track record of protecting against these traditional threats that are well understood.

Almost ten years on from the launch of 4G, 5G will be the first cellular generation to launch in the era of global cybercrime. This cybercrime activity is heavily funded by organized crime and nation sates. This is an era in which the exact same software that has contributed so much to driving the digital economy over the last ten years, is also routinely being weaponized to steal, expose, compromise or block access to data whether it is in use, at rest or in motion.

Most enterprises have many years' experience in evolving their data protection strategies while capturing the opportunities of virtualization and cloudification of IT infrastructure that's being driven by the global cloud service providers like Google, AWS and Microsoft. Yet, despite billions of dollars of investment in data security, it seems like barely a day goes by without a front page media headline pointing to a large scale corporate data breach, with large scale impacts on the business itself, its employees, customers, suppliers and investors.

So while enterprises certainly look to the 5G ecosystem of leading telecom operators, cloud providers, vendors and systems integrators to help understand the opportunities of 5G, they also expect guidance from that same ecosystem around how to understand and mitigate any new risks that the 5G architecture may pose to their data security posture.

Without a doubt, the 5G architecture does pose new risks to the security of enterprise data. Ironically many of these new risks take the form of the flip side – or downside – of the very same features of 5G which create so many of the new opportunities.

From the perspective of the telecom operator's 5G network, four new challenges are worth highlighting in particular:

4 Challenges for the Telecom Operator's 5G Network

i) The Transition to Network Function Virtualization (NFV)

With NFV, each network function no longer resides on its own proprietary hardware platform where this physical isolation provides a high level of protection. Instead it now resides in software as a Virtual Network Function (VNF) running on Virtual Machines (VMs), or Container Network Function (CNF) running on containers alongside other types of VNF /CNF, all sharing within a given tenant the same virtualized infrastructure. Left unprotected, there is a high risk of VNFs /CNFs interfering with one another or of cross-contamination of malware from one container to another, resulting in data leakage.

ii) Distribution of Network, Storage and Compute Functions to the Edge

To deliver the ultra-low latency required by many 5G applications, applications may require to be hosted out at the edge of the network. This is the 5G MEC (Multi-access Edge Computing). This layers additional challenges on top of the base NFV transition challenge noted above. For example, remote sites have less physical protections than centralized data centres. Cost or performance considerations may also require that collected data be stored at the edge rather than transported back to a more central location which already offers a high level of security. That locally stored data will need to be stored securely from both a physical and logical perspective, and the methods used for the encryption or the access to this data will have to cope with the ultra-low latency requirements.

iii) Resource Sharing Between the Telecom Operator and the Enterprise

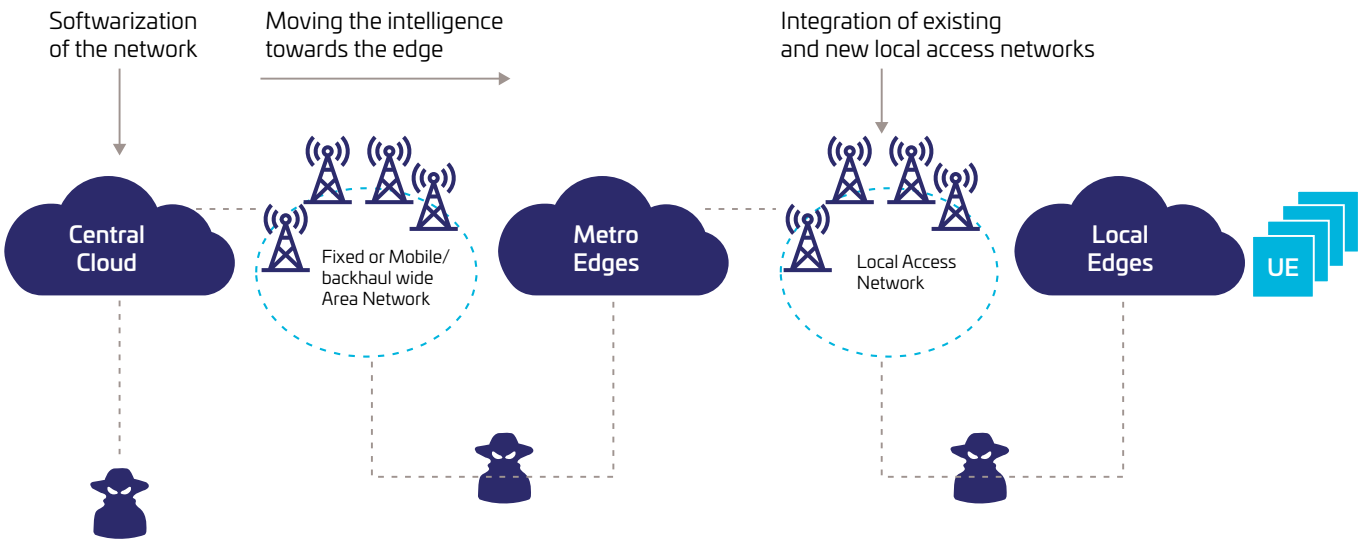
Some potentially popular IoT use cases assume the collection and storage of data at the edge of the network – as well as the application of analytics to that data at the edge. This creates a largely unprecedented scenario in which cost and space constraints may require that the telecom operators' network and the enterprise' own data and analytics software may share the same physical hardware or virtualized infrastructure. The same kind of isolation that is required within the telecom operator's own networking context will also be needed between the enterprise's own instances of analytics and the telecom operator's network. This requires a clear separation of duties between the enterprise and the operator and likewise the granularity of perimeters to manage will be extremely fine.

iv) Zero-Touch Automation

Zero touch automation is an intrinsic component of 5G networks. This is due to several factors, such as the usage of cloud-native virtualisation technologies, with Virtual Network Functions (NFVs) running on containers on top of a virtualized infrastructure, which have to be rapidly instantiated, scaled, updated or terminated. Due to the way that 5G and virtualisation enables applications and services to be spun up very quickly and at lower cost, zero-touch automation tends to be thought of primarily as an opportunity. The down side – which gets less attention – is that from a data protection standpoint, zero-touch automation propagates risky or even dangerous outcomes just as rapidly as it propagates beneficial ones. In a zero-touch environment, corrupted insights derived from compromised devices will automatically trigger the wrong events which will have negative or even severe consequences, depending on the use case. A high level of automation is also very helpful to the distribution of malware as well as to the propagation of leaked data. In the same way, a non-controlled automation could lead to the propagation to the entire network of a wrong decision in one small part of the network, which could have big effects on the service availability.

In the 5G era, many of the traditional assumptions around data security are no longer valid. Traditional security protections between the end device over the 5G New Radio (NR) remain necessary but are no longer sufficient. A lot of 5G traffic won't just be destined for the 5G core, but will also head for hosted analytics platforms that could just as easily be located a few meters away from the endpoint where it originated from or several miles away in a cloud. To command the trust of the enterprise, 5G security will need to combine underlying hop-by hop security such as built-in encryption on the 5G (NR) radio interface and integrity protection with end-to-end protection from the device to the corresponding application in the cloud.

The next chapter describes the solutions that need to be implemented to give enterprises the underlying trust they need in the 5G ecosystem. Subsequent chapters then go into ways in which IoT services, connectivity solutions and authentication capabilities can build on that platform to drive the 5G ecosystem and realize its huge potential.



Creating a Fabric of Trust Across The 5G Network

Due to the way enterprise data is embedded alongside the network functions at the core and the distributed edge of the 5G network, protections need to be built into the network and made available to enterprises to give them the foundation of trust they need to unlock high levels of investment in new 5G use cases. Traditional security controls applied to data in transit within the core will also need to be applied to data in transit between the edge clouds (for example an autonomous vehicle handing off between two remote edge nodes) as well as between the edge and the core.

5 Key Principles for Creating A Fabric of Trust in Network Slicing

1. Any given slice will need to be fully autonomous from a logical perspective, even if it shares the same underlying physical infrastructure.

2. Enterprises will need the option of requiring authentication for each and every one of their network slices. They will also need the option to authenticate onto one slice - which either they or a trusted third party like a telecom operator would host - followed by authorization onto other slices.

3. There needs to be a guarantee of isolation of each unique slice and each unique VNF within each slice. A VNF that is spun up in a given slice needs to function solely within that slice, with no way for it to go off-slice and gain access to another slice.

4. To be trusted as providers of hosting and analytics services as well as basic connectivity providers, telecom operators will need to demonstrate isolation of the VNFs within a commissioned slice from the enterprise's data and applications. This is to assure that they cannot view the enterprise's data. Enterprises will expect to see these competences demonstrated not just in the virtualized core, but also at the edge where distributed resource limitations may make this more challenging.

5. For maximum elasticity there will undoubtedly be demand for stateless VNFs. Those cases will necessarily require localized storage per slice at the edge of the network which will require protection.

Enterprises will call out these requirements to the telecom operators, cloud providers and vendors that are responsible for building this new foundation of trust. The rest of this section describes the key capabilities that these primary infrastructure providers will need to deploy in the network to meet those requirements.

Transparent data encryption with granular access control across multi-vendor VNFs/CNFs running in the network

One of the key components of a trusted 5G architecture is the integrity of the virtualized infrastructure and the confidentiality of the data flowing inside it. Cloud providers and their enterprise customers have been securing highly automated, virtualized infrastructure at scale in the cloud for many years – from securing the cloud provider's own infrastructure, to enterprises protecting sensitive workloads, and meeting compliance requirements such as GDPR. The telecom sector is making up ground now, with many mobile operators currently investing in these measures as they virtualize their existing 3G and 4G infrastructure. Measures for securing virtualized infrastructure that are in their infancy today will nevertheless need to reach maturity when 5G is rolled out. During this transition, while the MNOs focus on building the NFVI (Network Function Virtualization Infrastructure) leveraging general purpose hardware, the NEPs (Network Equipment Providers) are focusing on building VNF/CNF software instead of PNFs (Physical Network Functions) with specialized hardware tightly coupled with software.

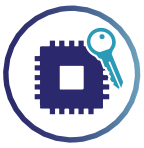
As the telco industry moves to leveraging virtualization, it implies that NFs are no longer isolated. In the past, with PNF and proprietary security implementations, any vulnerabilities led to limited impact as they could be easily isolated. This is no longer the case as VNFs/CNFs run on shared infrastructure. MNOs now need to take additional care to ensure that VNFs/CNFs are well segregated to minimize impact on the network. In parallel, 5G networks are expected to handle huge amounts of data as new services are deployed. This combination of virtualized infrastructure and increased amounts of data processed by 5G networks creates a strong need for data segregation. Leveraging data encryption with policy-driven, fine-grained access control can easily help address this challenge.

Furthermore, as MNOs will be deploying VNFs/CNFs from multiple vendors in their virtualized core infrastructure, relying on native data security solutions from VNF/CNF vendors leads to silos and gaps in security. It can also increase the likelihood of data breaches due to misconfigurations/human errors resulting from having to manage multiple disparate native security mechanisms. Deploying VNF/CNF vendor agnostic data encryption and access controls can greatly increase security while minimizing operational overhead of managing multiple systems.

It's important to highlight that to successfully adopt this approach, there are 3 key requirements: (i) Ability to "transparently" apply encryption to VNFs/CNFs, i.e. without requiring any modifications by VNF/CNF vendors, (ii) Ability to support both stateful and stateless VNFs/CNFs, (iii) Need for software based encryption solution. Hence, we find transparent file/volume level encryption to be the best solution. It brings the ability to "transparently" provide fine-grained access control to encrypted data while being flexible to handle stateless or stateful implementations across any kind of data store (databases, big data, file/volume level for VMs/containers with local/remote storage like DAS/SAN/NAS etc.). This minimizes impact on time to market while meeting security needs; making transparent, VNF/CNF vendor agnostic encryption of data with fine-grained access control the best guarantor of data segregation and confidentiality in MNO's virtualized infrastructure. The benefits of fine-grained access control along with transparent encryption are an important consideration with this solution as it addresses the need for data segregation while ensuring confidentiality, thus increasing the level of security as compared to that offered by storage-level encryption solutions (which mainly protect from physical storage device theft).

5 Key Capabilities to Create Trustworthy 5G Virtualized Networks

Transparent data encryption with granular access control across multi-vendor VNFs/CNFs running in the network



Enterprise controlled encryption at the edge

Key and encryption management solutions for centralized lifecycle management

Ultra-low latency encryption for "anyhaul" transport

Root of Trust for securing the infrastructure

Key and encryption management solutions for centralized lifecycle management

Encrypting sensitive data is a very important first step, but additional steps must be taken. In security parlance, that is the equivalent of locking your front door but then leaving the key under the mat. Centralized lifecycle management of the encryption keys is a critical accompaniment to data encryption. The full lifecycle of a given encryption key needs to be managed from the moment that it is first generated and registered, through storage distribution and rotation, until its eventual destruction. This lifecycle of the key needs to be managed according to the specific policy that is chosen, applied and enforced by the organization. The management needs to be centralized in a secure manner to ensure the protection and control of the data at rest and in use. In addition to managing the life cycle of encryption keys, management solutions should also be able to manage other data confidentiality protection functions, such as encryption operations on centrally or remotely located databases, files and folders, storage devices and/or at the application level. This flexibility to handle encryption at any layer within the overall system would help address specific data security needs depending on the data classification level while centralizing the associated key lifecycle management capabilities.

Another important aspect to consider with centralized key and encryption management solutions is the reporting capabilities built into the solution. With the explosion of data to be managed in 5G networks, visibility on sensitive data access becomes essential to monitor the security of the networks and meet regulatory needs. Besides generating logs, it's important to be able to integrate/feed the logs into SIEM systems for alerting and remediation. Key benefits of such a solution deployed in the telco's core infrastructure are:

1) Identify unusual sensitive data access patterns and attacks on the centralized key and encryption management solution from unauthorized users/processes

2) Accelerate detection of insider threats and APTs (Advanced Persistent Threats) by:

a. Monitoring process access to protected data for anomalous patterns of use that could indicate a process has been coopted by malware

b. Pinpointing unusual patterns of user access to protected data that indicate malware (or a malicious internal user) could be stealing data

3) Create audit and compliance reports to meet regulatory and internal security requirements

Root of Trust for Securing the Infrastructure

At the heart of a trusted computing or networking environment there should always be a root of trust (RoT) - a set of functions that are inherently trusted by the operating system or systems creating a foundation of trust. An HSM is the most reliable anchor of trust in a digital world. It is a dedicated hardware system, optimized for cryptographic processing, that physically and logically secures the cryptographic keys and cryptographic processing which enable the nodes or endpoints in any network to trust the integrity of the data that is being presented to them.

Whereas the direction of the telecom and IT industries appears to run in favor of the primacy of software, most telcos already use HSM appliances as a RoT to secure PKI infrastructure (for e.g. for securing communication between base stations/small cells and the backend in 4G mobile networks) and code signing operations. HSM usage is now expected to expand to provide RoT to PKI infrastructure related to NFVI deployment requiring certificates for OpenStack, Kubernetes, CI/CD channels, VNF/CNF signing, other apps/open source tools, service mesh etc.

Some telcos also use HSMs in their authentication centres due to the high-assurance, FIPS-validated and Common Criteria-certified security they provide. With the possibility of enterprises leveraging 5G network slices for mobile private networks, it's expected that cryptographic authentication processing secured by HSMs will grow in demand. Another use case that could benefit from the security offered by HSMs is related to a new subscriber privacy mechanism that has been introduced in 5G, i.e. SUPI/SUCI (Subscription Permanent Identified/ Subscription Concealed Identifier). HSMs could be used to secure the key generation, storage, and cryptographic processing in the core network to ensure utmost security for subscriber privacy during the 5G authentication process.

It is also noteworthy that the majority of the big cloud providers continue to recognize the value that a hardware-based root of trust can offer to businesses that want to be leaders in data protection. Cloud providers such as Microsoft Azure and IBM all offer HSM resources to their customers in the form of dedicated hardware products deployed on-premises, as does Thales with Data-Protection On Demand which is operated in an "as a service" model and is cloud agnostic. Those who are poised to lay the foundations of the 5G ecosystem need to ask themselves whether they too need to follow the same path or whether the available software-based alternatives without a hardware RoT for securing the most sensitive cryptographic operations in their core network will suffice. The choice is quite stark: the root of trust provided by an HSM creates the anchor for the most secure possible chain of trust at the heart of the 5G ecosystem.

Enterprise-controlled data security at the edge

There are important new challenges at remote sites where data is also being analysed and manipulated at the edge of the network. Edge cloud service deployments are expected to grow to drive new services requiring ultra-low latency and high throughput. These new services are expected to address verticals like healthcare, manufacturing, retail and media/entertainment etc.

With more and more data expected to be processed at the edge, the means to secure sensitive data needs to be planned for. The concept of a shared responsibility model between cloud service providers and enterprises should be extended to these Telco edge cloud sites. This implies that the data managed by enterprises leveraging Telco Edge Cloud is the enterprise's responsibility. Enterprises can secure data processed in edge sites by leveraging encryption along with fine-grained access control mechanisms. There are broadly 2 options available for encrypting data in such scenarios: BYOE (Bring Your Own Encryption) or native encryption from the edge cloud platform provider.

BYOE enables enterprise customers to leverage their existing encryption and key management solutions and extend it to the infrastructure provider at the edge. It helps address aspects like Separation of Duty between storage infrastructure provider and encryption provider, enabling the highest control by the enterprise on sensitive data that it owns/manages.

In recent years, public cloud service providers have enhanced their native encryption capabilities by giving enterprises greater control on the cryptographic keys by enabling mechanisms like BYOK (Bring Your Own Key) and HYOK (Hold Your Own Key). Depending on how the Telcos provide edge cloud capabilities to their enterprise customers, they would need to either build similar capabilities for their Telco edge cloud offerings or they could let their enterprise customers leverage such capabilities from their public cloud provider partners (this applies only to scenarios where there is a partnership between Telco and Cloud Service Provider for edge cloud offering). Compliance to regulatory requirements will also drive the need to protect sensitive data like: PII, PHI, PCI etc. that are collected and processed at the edge for new services addressing healthcare, surveillance, smart manufacturing, entertainment and media.

Ultra-low latency encryption for 'anyhaul' transport

As well as in use and at rest, data also needs protecting in transit. Just as with 4G, 5G will provide in-built encryption on the 5G NR radio interface; but from the cell site back across the network there won't be any encryption baked in.

Instead, as in 4G, 5G operators will need to add their own encryption across the backhaul, fronthaul even the midhaul (Interconnect between remote edge clouds) networks, sometimes referred to as "anyhaul".

In 4G, the encryption standard typically used in this context is IPsec. While well-suited to 4G, IPsec is much less suited to 5G. In particular, because of the overhead it carries, IPSec will jeopardize 5G's ultra-low latency requirements.

Other high speed, ethernet encryption solutions operating at 100 Gbit/s and adding nothing more than a few microseconds of latency are already widely available on the market and need to be considered as an alternative to IPsec for 5G.

The next chapter goes into ways in which IoT service lifecycle management must evolve to meet new requirements in the 5G ecosystem.

Only as secure as the weakest link

All of the above key components of a trusted 5G fabric are complementary and inter-linked. Enterprises that handle sensitive data as part of their core operations should consider them all to be mandatory on the basis that the overall security of the trusted 5G fabric can only ever be a strong as its weakest link.

2. How to provide trustworthy 5G IoT device lifecycle management

In discussions around the vulnerability of the global digital economy to insecure IoT devices, much of the attention has focused on the consumer market and the way botnets are able to leverage insecure household 'things' to carry out attacks like the Mirai botnet attacks of October 2016.

Given the way the network is evolving towards 5G, however, protecting enterprise-generated IoT data will be just as challenging. This section looks at the risks to enterprise IoT data and how, for many organizations, outsourcing end-to-end IoT device life cycle management will drive their trust in the 5G ecosystem.

The traditional M2M business model of a large industry vertical investing its own resources in assembling partnerships to seed its own cellular devices with authentication credentials, and then deploying them into a closed, centralized, inflexible architecture using only the company's own servers and its own analytics, is starting to break down.

In the 5G ecosystem, organizations will expect to be able to dynamically change both network and hosting providers whenever it is advantageous to do so. If they can trust the ecosystem, then they would also like to be able to leverage third party data and analytics to add additional value to the insights they derive from their own data. This requires extending the trusted fabric into these domains. Enterprises need to retain strict control of their data. They also need to know which third party data and analytics resources can and cannot be trusted in this more open data-brokering environment.

Global cloud providers such as Microsoft, IBM and AWS are lending their considerable weight to this trajectory on the cusp of the 5G era. In 2017, for example, AWS launched Greengrass, its IoT proposition for distributed environments.

This gives developers the tools they need to be able to code directly into distributed edge environments with local compute, storage and analytics resources. This is for use cases that are deployed many, many miles from traditional centralized clouds. 5G complements this with the distributed characteristics, ultra-low latency, choice of optimal connectivity, and network slicing that 4G can't provide.

Nokia is one of several companies partnering with AWS to drive adoption of Greengrass at the edge of the 5G network.

In the evolution from M2M, IoT in 5G has 3 new characteristics:

A. A variety of devices with greatly varying capabilities, most of which will not use cellular connectivity.

B. Very different types of connectivity with very different performance characteristics to support applications that have very different requirements.

C. A very much more open, flexible and distributed architecture.

5G has four main impacts on IoT devices:

1. The time that some devices will need to be in the field is lengthening – to twenty years or more in the case of some narrowband IoT connected objects. Ways need to be found to ensure these devices and applications are updated throughout their lifetime with the latest firmware and software.

2. Increasingly the connection of these objects is via shared or public infrastructure and protocols. In this environment there is a need to protect both servers and clients from unauthorized access. Identity credentials need to be not just provided but managed as well.

3. The diversity of data and diversity in the use of the network platform may create conflicts within the device or the connectivity that may lead to service disruption.

4. Ideas for new IoT deployments will increasingly be application-driven, by developers with little or no experience in the assembling of all the components of an end-to-end solution themselves in what is a far more complex environment than the traditional M2M model.

IoT device lifecycle management entails provisioning or seeding the security credentials and the connectivity capabilities into the device at the point of deployment. It assures that the choice of connectivity at the point of deployment meets the requirements of the application. It also ensures that the device consistently presents itself as a trusted device to the app in the cloud.

Building advanced IoT solutions that are robust and secure doesn't just require resources and capabilities at the deployment stage. Long before you get to that point, a lot of resources have to be invested in initially qualifying the business model, proto-typing, as well as the deployment and testing of the prototype. Just as they have always been able to, dating back to the early M2M days, some of the largest organizations have enough resources of their own to navigate and implement the new device-to-cloud data security challenges of their IoT deployments in the 5G context.

Most organizations, however, do not and therefore most face a stark choice. They can either deploy IoT solutions into the exciting but risk-laden 5G environment with inadequate trust mechanisms built in – and leave themselves very vulnerable to data breaches. Alternatively, they can dramatically reduce that risk by outsourcing these requirements – for the pre-commercial as well as the commercial phases – to a specialized provider of IoT lifecycle management services. This provider should have expertise in zero-touch device provisioning of data protection and connectivity credentials, together with end-to-end data protection solutions in the device, the network and the cloud.

3 Key Capabilities for 5G IoT Lifecycle Management

I. End-to-end lifecycle management of an IoT device's keys and certificates.

II. Secure management of updates of both the firmware on the device and the application. This ensures that the updates that are received are not only trusted but also adhere to the unique performance requirements of the specific application.

III. Monitoring for any unauthorized behaviour on the connectivity bearer or by the device and troubleshooting to remediate issues.

3. How wireless modules can address new 5G IoT use cases

Due to the accelerating market momentum behind IoT, wireless modules can be expected to play a bigger role in the 5G ecosystem than in the case of previous generations of cellular – whether an enterprise manages these modules itself or whether the management is outsourced to a third party. This section details how some of the requirements of wireless modules change significantly.

5G modules will need to offer the same seamless integration with the 5G NR as 4G modules support with LTE. Modules will also be manufactured centrally and then need to support the same on-demand connectivity so that they can be taken out of the box and deployed instantly anywhere in the world. The provisioning of the specific operator profiles and security credentials of any one from hundreds of mobile operators around the world will be substantially the same as in 4G.

The need for a broad range of modules to support multiple frequency bands will remain. If anything, there will need to be even more to support all of the additional spectrum bands that are eligible to use 5G.

Besides these traditional table stakes, however, 5G modules will have to support some very different characteristics to previous generations of modules. 5G will see enhanced modules defined according to three broad categories:

- Machine Type Communications;
- High Speed Communications
- Ultra-Reliable Communications

Of the three, Machine Type Communications (MTC) modules will resemble their 4G counterparts most closely. Changes in requirements here will come mostly in the case of those MTC products that are optimized for Massive IoT – namely long battery life and very high coverage to serve the Lower Power Wide Area (LPWA) market. A smarter device-centric architecture can be envisaged in 5G with on-board Quality of Service measurement and additional features such as device-to-device communications will be leveraged. New security measures will also be required to support non-IP communications protocols like the new NB IoT NB1 and NB2 protocols in the context of devices being deployed in the field for ten or more years.

Modules optimized for High Speed Communications (HSC) together with low latency are an entirely new category that are only viable with 5G. Anyone who questions the market's real appetite for 5G modules – rather than making do with 4G products – should take a look at the automotive sector.

Automotive leaders are aggressively driving demand for 5G modules. Via industry associations like the 5G Automotive Association (5GAA), which includes Audi AG, BMW Group and Daimler AG among its eight founding members, leading automotive companies have already put several years of R&D into readying their new product lines to be 5G-ready when the first networks are launched.

Some ecosystem partners have already arrived at a first product definition for a 5G high speed module for automotive. This is for use at frequencies below 6 GHz, for coverage reasons. Detailed specification work on the Ultra Reliable Communications (URC) category has yet to get underway in 3GPP and 5GAA. These Ultra-Reliable products can therefore be expected to come to market after the combined High-Speed LTE with "early-drop" 5G New Radio products which are mainly aimed for consumer use and will have a short lifetime (up to 2 years).

First fully standardized standalone 5G automotive products will likely launch in the late 2021 to early 2022 timeframe. This is due to the lifetime and performance level requirements from the automotive industry. A car can have a lifecycle of at least 10 years and therefore the "early-drop", "Non Standalone" implementation, where 5G is not implemented in the core network and connectivity is relying on the existing LTE core, will not be sufficient for this segment. Product development will likely draw upon many of the same skillsets required for the High-Speed Communications (HSC) module market in terms of use case definition, prototype development testing and deployment.

At a more general level, because of the chain of trust requirements relating to data protection already outlined in this paper, 5G modules will have to interact with the 5G network in a fundamentally different way to previous generations of wireless modules. They will need to support data protection and other security features that go way beyond just traditional compliance to 3GPP security standards. They will need to offer service providers and their enterprise customers an extensive range of security features so that they can navigate the unique requirements of the applications they support. That includes specific requirements for authentication and authorization onto a network slice or authentication credentials in relation to other devices and instances of compute, storage, network and analytics in the network or in the cloud.

The next chapter examines how the end device must be securely authenticated for access to the 5G system.

4. Which value the SIM brings in 5G

This section focuses on the role of the SIM in 5G including some new features that broaden the possibilities of usage of the SIM.

The starting point is all about security. Like over 2G, 3G or 4G networks relies in the fact that the SIM carries the symmetric key used for the authentication over a 5G mobile public network. But this fundamental aspect is only one facet of the additional security features that can be leveraged with the 5G SIM. For instance, for a given authentication, an MNO could be willing to have the possibility to choose between several authentication algorithms, all being stored within the 5G SIM. Or it could request to have ways to update the symmetric key stored within the SIM tamper-resistant, certifiable, dedicated hardware component. This is made possible with 5G SIMs.

But these new ways of handling the authentication are opening doors to other fields of usage. Keys 5G features like slicing or new promising areas like private networks, the so-called non-public networks, can as well leverage the SIM's unique capabilities to provide a secure authentication. A service provider could simply be willing to rely on the SIM to authenticate on a non-public network, just because they would not accept any compromise with security.

The second step is about privacy. In our societies, end-user are more and more caring about the way their data is used. In parallel with this, we have seen all around the world more and more privacy regulations being published, like the European GDPR. One of the fundamental new point brought by 5G is all about this. Over 5G, the user identity, called the SUPI (Subscription Permanent Identifier) is not circulating in clear like the IMSI over 2G, 3G or even 4G. It is now encrypted or to say it "more 5G", "concealed" and becomes then the SUCI (Subscription Concealed Identifier). The way this is done relies entirely on the SIM, which is the entity storing the keys and parameters required for this. Without a SIM with the proper capabilities, there is no new 5G privacy capabilities for the end-user.

The third step is about quality of service. For sure, the 5G SIMs take into consideration the fact that 5G parameter formats such as the location information are introduced. But it goes way beyond: there are now new possible ways to perform Steering of Roaming (SoR). The goal is to solve efficiency issues: in 3G/4G, SoR leads namely to either a poor user experience (the user may need several minutes before they can connect on a visited network and may experience lack of coverage in the visited country) or from ineffectiveness since the SoR request may be subject to anti-steering or simply be filtered by the visited network for security reasons. Leveraging 5G SIM, and 5G OTA notably to provide the requested confidentiality, MNOs can benefit from the new 5G SoR defined in 3GPP to remove all these restrictions: users have immediate access to connectivity when roaming, no lack of coverage, no anti-steering or SoR requests being filtered.

At the end, there are as well new possibilities to store within the SIM application-related parameters that can dramatically enhance the end user experience. This is called the URSP (User Equipment Route Selection Policy) and covers parameters like the DNN (Data Network Name, the equivalent of APN over 5G), the slice reference (NSSAI – Network Slice Selection Assistance Information), the type of network access (3GPP or non 3GPP) etc. All of these are designed to give to an application the best route over 5G networks.

These novelties will work alongside technologies and form factors that are not specific to 5G, and will largely take advantage of it. For instance, the eSIM or "embedded SIM" is a SIM with the capability to be remotely (Over-The-Air) personalized, in-line with the specifications published by GSMA. These eSIMs have been used by a large variety of devices, from smart watches to smartphones, and from cars to tablets in the last few years. These eSIMs have a nice promise as their unique capability matches with the need for more flexibility, and ubiquity in the delivery of the device connectivity. Although they have been initially deployed over 4G networks, the eSIMs can also be 5G enabled and will for sure accompany the rise of 5G over the next years, burgeoning into new devices and being a key enabler of new use cases.

This section discusses how the format, function and management of the network connectivity credentials in end devices has to transform in order to align with - and fully exploit - the new capabilities embedded in the 5G network. In particular it addresses the evolution from the distribution of one set of user credentials in a physical SIM card to the distribution of multiple sets of credentials that are remotely programmable in software.

Since the launch of the first secondary "companion" devices equipped with an eSIM - (3G Samsung Galaxy Gear S2 in 2016 followed by the 4G Samsung Galaxy Gear S3 and the Apple Watch Series 3 in 2017) numerous eSIM compliant devices including primary ones (e.g. smartphones) have been introduced. By the end of 2018 Apple designed all its new iPhones to be XR and XS eSIM compliant and even eSIM-only ones, meaning there was no longer a SIM-tray, like the Motorola Razr in 2019. This marked a new era in mobile communications in that this usage of SIM was the beginning of a transformation from the physical to the digital management of end device security credentials.

The eSIM reduces the cost associated with the manufacturing, shipping and channel logistics of discrete physical goods and enables the end users' digital journey. Cellular-connected devices can be remotely personalized, receiving the equivalent of the SIM card's content (Connectivity Provider's Profile) over the air. This capability facilitates bringing connectivity to new types of devices, as for instance mentioned above for the secondary devices, and offering new possibilities to end-users as it virtually enables the management of several SIMs in one device and dynamically swapping between them.

5G will launch just when the new eSIM market is starting to scale up, making it the first generation of cellular to launch in a buoyant eSIM market. For 5G, whether it's deployed in a 5G consumer smartphone or in a 5G IoT device, eSIM will be a key catalyst. In addition to that, a lot more is required of a 5G eSIM.

Advanced connectivity

- A discovery service so that users can choose their own basic connectivity subscription.
- A more defined mechanism for seamless roaming between 3GPP and non-3GPP wireless access technologies.
- For enterprise or content services like Netflix, it is likely that the subscription including the connectivity will be paid for (and bundled) by the content provider. The 5G eSIM should enable this to work seamlessly alongside other subscriptions for content and/or connectivity.

Network slicing authentication and security

- Enabling enterprises to leverage their credentials to preselect network slices.
- To support devices with access to multiple network slices, a 5G eSIM needs to support multiple sets of authentication and authorization credentials. Solutions need to comply with the GSMA's RSP specification and be made available per slice to allow credentials to be downloaded for each slice to which the device has access.
- In a multi-slice device use case - where a device is connected to both an enterprise slice containing the enterprise's sensitive data and a commercial or consumer slice - the eSIM and other trusted elements in the device need to enforce device-level isolation. This is to ensure that malware cannot be inserted to leak data from the enterprise confidential slice to the consumer slice where a malicious entity is spying and potentially exposing the data from one slice to the other. Undoubtedly the new role of the 5G eSIM is in enabling the fabric of trust for network slicing.

In both the enterprise and consumers contexts, the flexibility and range of options available when eSIMs are used to access the 5G network will open the door to a cockpit of subscription management so that users can see what devices they have, what each are consuming and what they're paying.

In the consumer context, it is increasingly common for operators to offer single data 'bucket's or family plans which users can draw upon using multiple devices. This will accelerate in 5G as the multi-access 5G network gives a sense of 'limitless capacity' with widely differing performance characteristics. This will also drive pricing away from volume-based billing towards more value-based billing models.

From a telecom operator perspective, enabling these single, multi-device contracts with variable billing models will require new capabilities on the network and device side. Upgrades to their BSS environments will be needed to manage credentials on the network side.

On the end device side, a new generation of subscription management platforms will be needed. These will be able to define the security and other features of a given network slice and then allow credentials to be seamlessly downloaded on a per slice basis in a few clicks.

In addition to being standards-compliant, subscription management platforms will need to enable easy integration with a telecom operator's legacy subscriber physical SIM provisioning environment, including its current reporting, monitoring and other operational aspects. As telecom operators move from purchasing physical SIMs to buying eSIMs as digital subscriptions, management platforms will need to support that transition according to formats that telecom operators aren't currently familiar with.

5. How software licensing can help in a fully virtualized environment

Another key requirement for the 5G ecosystem to flourish is a software licensing regime that provides network operators with the flexibility they need to take full advantage of virtualization with automation and orchestration.

In order to reduce their cost base, telecom operators are no longer prepared to accept a one-size-fits-all perpetual licensing model whereby they pay for the entire software licence up-front for a vendor's VNF software for the right to use it indefinitely.

Operators increasingly expect to be able to opt for a variety of subscription models (where software is effectively leased for a monthly fee) or pay as you go models. These more flexible models align much more closely with the day-today reality of highly variable usage of different networking applications in a dynamic, software-driven, 5G network.

Operator demand is driving new dynamics in the telecom software market. Pure-play Independent Software Vendors (ISVs) are entering the market with new offerings. These are designed from scratch to align with the cloud-native requirements of the operators, unencumbered by the constraints of legacy hardware dependencies and legacy code. A number of the larger operators are also starting to develop their own software.

To successfully navigate this transition, software vendors need to up their game in terms of the more flexible purchasing models they offer but also in terms of the way they audit both the initial sale and subsequent consumption of VNFs to their telecom operators customers.

Vendors need licensing management capabilities to protect their software against intentional or unintentional abuse by their own customers. That means protecting their IPR in terms of protection against reproduction and reverse engineering. It also means accurately monitoring and tracking granular consumption of VNFs by customers to ensure that they are paying for everything they use. By allowing customers to self-monitor their software consumption, some start-up ISVs are particularly exposed to fraud.

Rather than having no software licensing management at all, many medium and large vendors suffer from the opposite vulnerability. As a result of having made many acquisitions, a lot of established vendors have inherited multiple licensing management systems across their portfolio.

Moreover, it is fair to say that many of these implementations are rudimentary. Many large and medium sized vendors therefore need advanced license management systems that can serve as a best in class platform in its own right. Critically, these should also serve as a single point of integration and unification for multiple disparate systems. Such disparate systems are holding many vendors back not only from meeting customer requirements – but from monetizing the value of their own products as effectively as possible.

About Thales

The people you rely on to protect your privacy rely on Thales to protect their data. When it comes to data security, organizations are faced with an increasing number of decisive moments. Whether the moment is building an encryption strategy, moving to the cloud, or meeting compliance mandates, you can rely on Thales to secure your digital transformation.

Decisive technology for decisive moments.

THALES

Contact us

For all office locations and contact information,
please visit cpl.thalesgroup.com/contact-us

> cpl.thalesgroup.com <