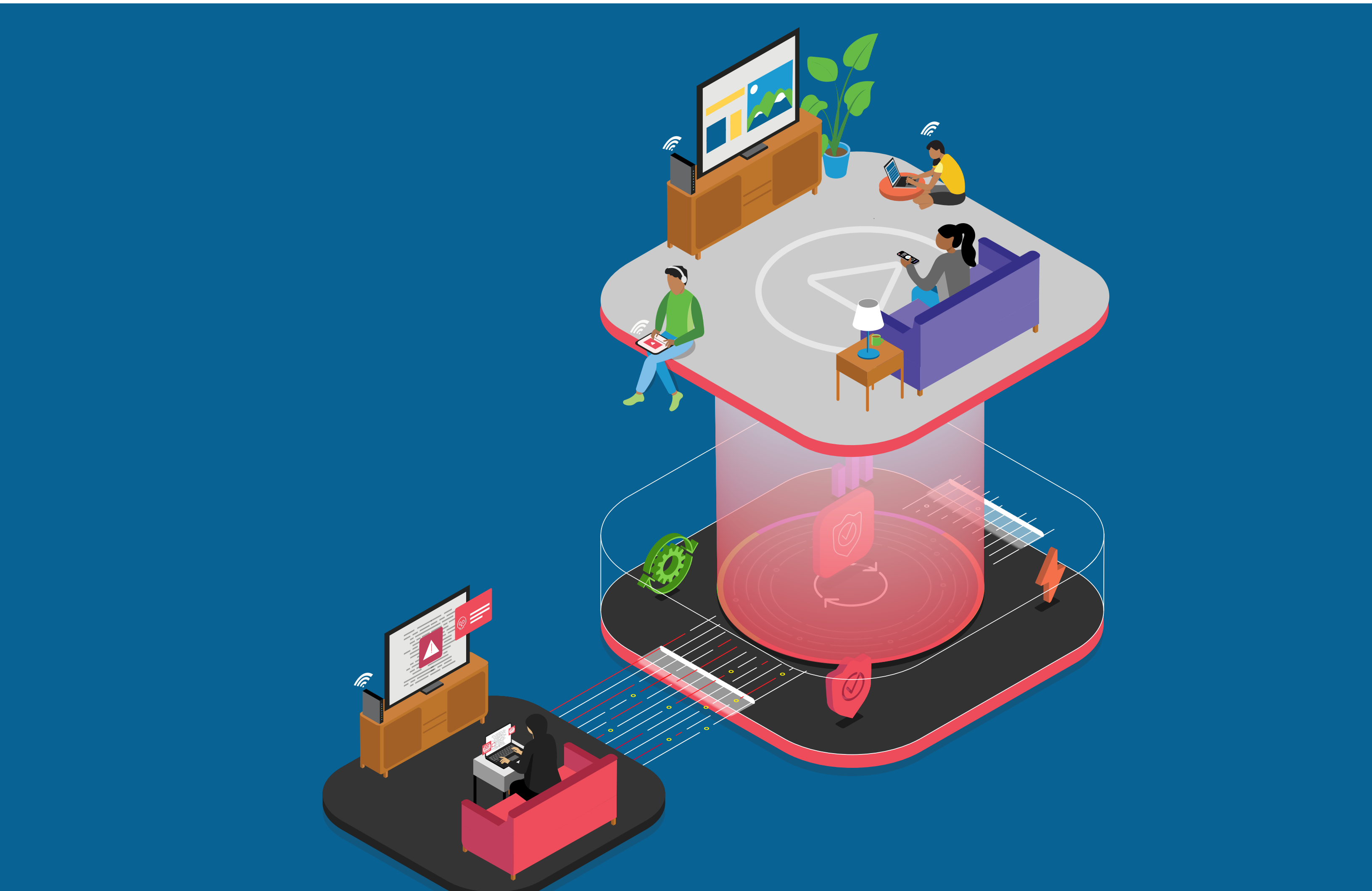
WHITE PAPER

Building Better 5G Security

How to harness and adapt IT best practices to protect 5G core networks



IN THIS WHITE PAPER, WE PROVIDE AN OVERVIEW OF THE KEY SECURITY CONSIDERATIONS IN EACH OF THE FOUR PLANES THAT MAKE UP A 5G CORE NETWORK.

As Mobile Network Operators (MNOs) implement standalone 5G core networks, they will implement service-based architectures that use IT concepts and protocols like HTTP/2, APIs, and microservices, relying on data centers at the edge of the network to support responsive connectivity.

This will give MNOs much greater flexibility and scalability than they have had in the past, giving them the versatility they need to support many more use cases and applications, and opening up new business opportunities. 5G also will extend the threat landscape and the attack surface that operators need to defend.

In this white paper, we provide an overview of the key security considerations in each of the four planes that make up a 5G core network:

- In the **data plane**, the N6 Interface, which sits between the UPF (user plane function) and the Internet, needs to employ an array of security tools, such as a N6 firewall, carrier-grade network address translation, protection against distributed denial-of-service (DDoS) attacks, domain name system security, and, increasingly, dedicated IoT firewalls. Pioneering 5G operators have found that employing a unified platform encompassing all of these security tools is more efficient and effective than sourcing them from individual vendors. Similarly, deploying DDoS mitigation capabilities on a SmartNIC (network interface card with an integrated field-programmable gate array) installed on a standard x86 server can be a cost-effective way to fend off attacks designed to overwhelm the computing resources in the data plane.
- In the **control plane**, the IT-based microservices architecture needs to be adapted to meet the needs of telecoms operators and to ensure that the various interfaces are secure. In particular, the service mesh between microservices needs to support security and traffic management for HTTP/2, PFCP, Diameter & GTP to enable interworking between 4G and 5G networks, while the security edge protection proxy (SEPP) requires JSON object signing. Additionally, a telco service mesh needs to allow for packet capture for lawful intercept purposes, as well as supporting traffic routing and overload protection.
- In the **management plane**, MNOs are increasingly looking to generate revenues by exposing APIs to external partners so they can program a slice of the 5G network for a specific purpose. Because this approach opens up the management plane to the outside world, it is vital that these APIs are properly secured.

- The **application infrastructure** needs to be protected by identity and access management systems and an anti-bot enabled web application firewall; microservices applications will also need API security. These security protections will need to be cloud and platform-agnostic so they can remain in place if the MNO decides to relocate an application from one environment to another. Relying on security measures tied to each cloud environment would result in inconsistent levels of protection and a degree of lock-in to that environment.

MNOs need to rigorously ensure that their 5G core network's service-based architecture fully accounts for the challenges inherent in operating a telecoms network that is broadly exposed to the outside world and handles large volumes of traffic from many different sources. However, with a considered approach and the right partners, it is feasible to secure all four planes of a 5G core network in an effective and cost-efficient way.

Drawing on F5's extensive expertise in IT and telecoms, we will explore how to combine best practices from the IT and telecom worlds to help secure the four planes of a 5G network.

The Fusion of IT and Telecoms

CARRIERS ARE TRANSFORMING FROM NETWORK OPERATORS TO IT COMPANIES AND 5G WILL ACCELERATE THIS MOVEMENT.
Christoph Aeschlimann,
CTO at Swisscom

As they roll out 5G networks, MNOs are evolving into IT companies. The progression from 2G to 3G to 4G involved transitions from one mobile architecture to another. The 5G transition is different. 5G networks make extensive use of IT concepts and protocols, such as HTTP/2, APIs, and microservices. In some cases, web companies, such as Rakuten, are using cloud-native technologies to become MNOs. They regard mobility as a new use case for their existing web scale architecture.

"Carriers are transforming from network operators to IT companies and 5G will accelerate this movement," noted Christoph Aeschlimann, CTO at Swisscom, at the 5G World event in September 2020. This shift is changing the technological landscape in a way that has major ramifications for how MNOs are organized and how they think about security. At the same event, Scott Petty, CTO of Vodafone UK, said: "The skills and cultural transformation is massive, and we will look like an entirely different company five years from now."

Figure 1 shows how IT underpins the service-based architecture of a 5G network. Each network function is instantiated as a microservice in a containerized or virtualized environment, instead of relying on dedicated physical boxes that provide a specific function. This service-based architecture, which is largely inherited from the IT industry, takes an API-centric approach to interactions, using HTTP as the protocol of choice.

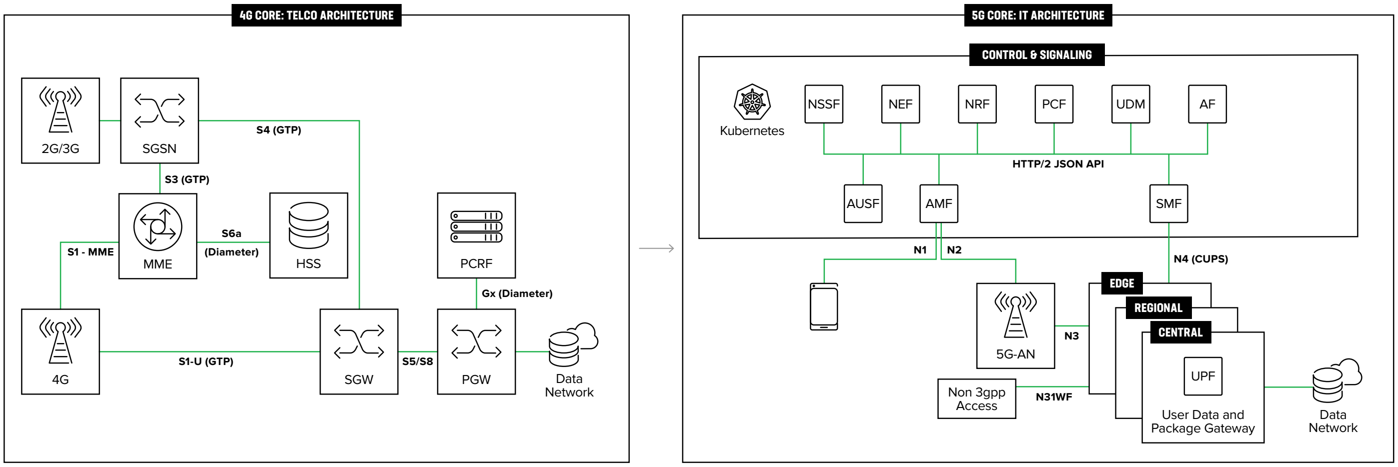


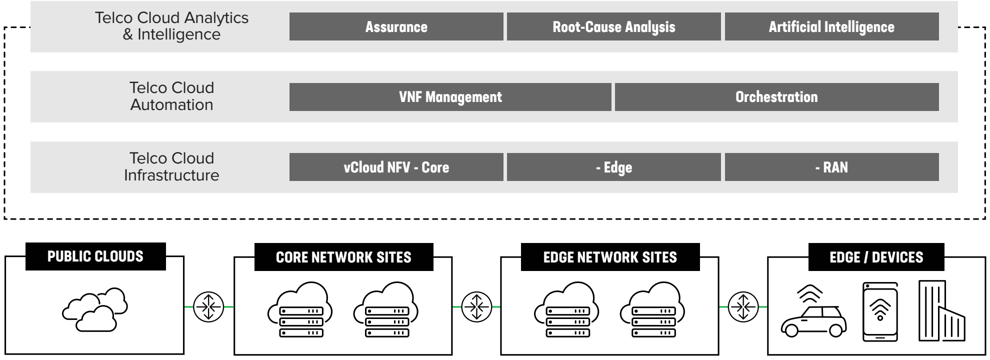
Figure 1: In 5G, the traditional telco architecture has been superseded by an IT architecture with the introduction of service-based architecture (SBA)

Another fundamental change brought about by 5G is the deployment of computing resources at the edge of the network. Many of the use cases for 5G, such as virtual reality, augmented reality and remote control services, rely on very responsive connectivity. To keep network latency to a minimum, these services will be delivered from data centers located relatively close to the end user. An application in a centralized public cloud won't be able to meet the latency requirements.

For example, in an industrial plant, the machinery may be generating massive amounts of usage and performance data, requiring near real-time analysis by an application located either on-site or in a MNO's edge data center. Using applications that reside inside the cellular network avoids the latency of round-trips to conventional public cloud infrastructures.

Many mobile operators have the ambition to run telco cloud architectures in which both network functions and applications coexist on the same infrastructure. 5G networks will have a multi-cloud approach encompassing central, edge and far-edge data centers (see Figure 2). Both the 5G network and the applications it delivers will be containerized and embedded in the same architecture.

Figure 2: A telco cloud will have data centers in the core and at the edge of the network



Implications for 5G Security

As 5G supports a much wider range of use cases than its predecessors, the threat landscape becomes much broader and the attack surface much larger. There are many different ways in which 5G networks could be attacked (see Figure 3). In particular, the virtualized mobile network components and the separate network slices being created for 5G use cases can open up entirely new security threat vectors. With a 5G core, an operator can give a customer access to a slice of connectivity, which can be tailored to the requirements of the use case. Customizing the network for third parties in this way requires new interfaces, which need to be secured. Each network slice also needs to be isolated from other network slices.

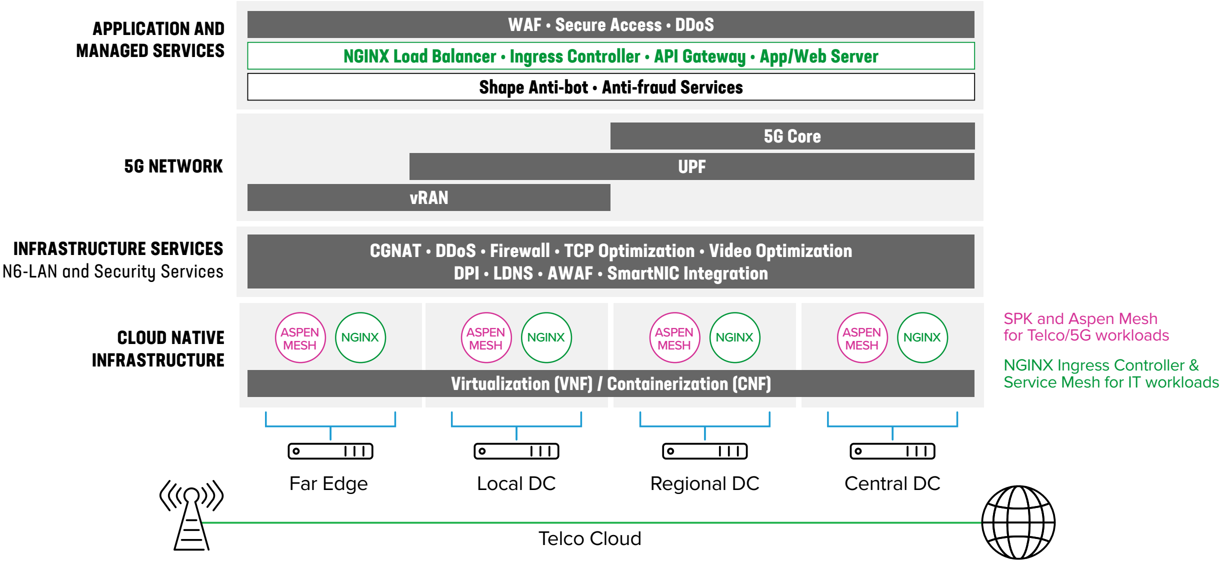

APPLICATION AND MANAGED SERVICES
WAF • Secure Access • DDoS
NGINX Load Balancer • Ingress Controller • API Gateway • App/Web Server
Shape Anti-bot • Anti-fraud Services
5G NETWORK
5G Core
UPF
vRAN
INFRASTRUCTURE SERVICES
N6-LAN and Security Services
CGNAT • DDoS • Firewall • TCP Optimization • Video Optimization
DPI • LDNS • AWAF • SmartNIC Integration
CLOUD NATIVE INFRASTRUCTURE
ASPEN MESH
NGINX
ASPEN MESH
NGINX
ASPEN MESH
NGINX
ASPEN MESH
NGINX
Virtualization (VNF) / Containerization (CNF)
SPK and Aspen Mesh for Telco/5G workloads
NGINX Ingress Controller & Service Mesh for IT workloads
Far Edge
Local DC
Regional DC
Central DC
Telco Cloud

Figure 3: The flexibility of 5G networks could also make them more vulnerable

The telco architecture that runs a 5G network can be divided into a data plane (also known as the user plane), the control and signaling plane, the management plane, and the application infrastructure. Each of these planes is vulnerable from both the inside and the outside (see Figure 4).

- The data plane, which carries the network user traffic, is connected to the Internet and the world of external applications, making it vulnerable to attack.
- The control plane is connecting or interconnecting with roaming partners, introducing an element of risk.
- In the management plane, MNOs will expose APIs to external MNO partners to make use of, or to even program, some aspects of their 5G network for specific purposes, such as IoT use cases that employ network slices.
- In the application infrastructure, third parties can deploy their applications inside the network, potentially introducing vulnerabilities.

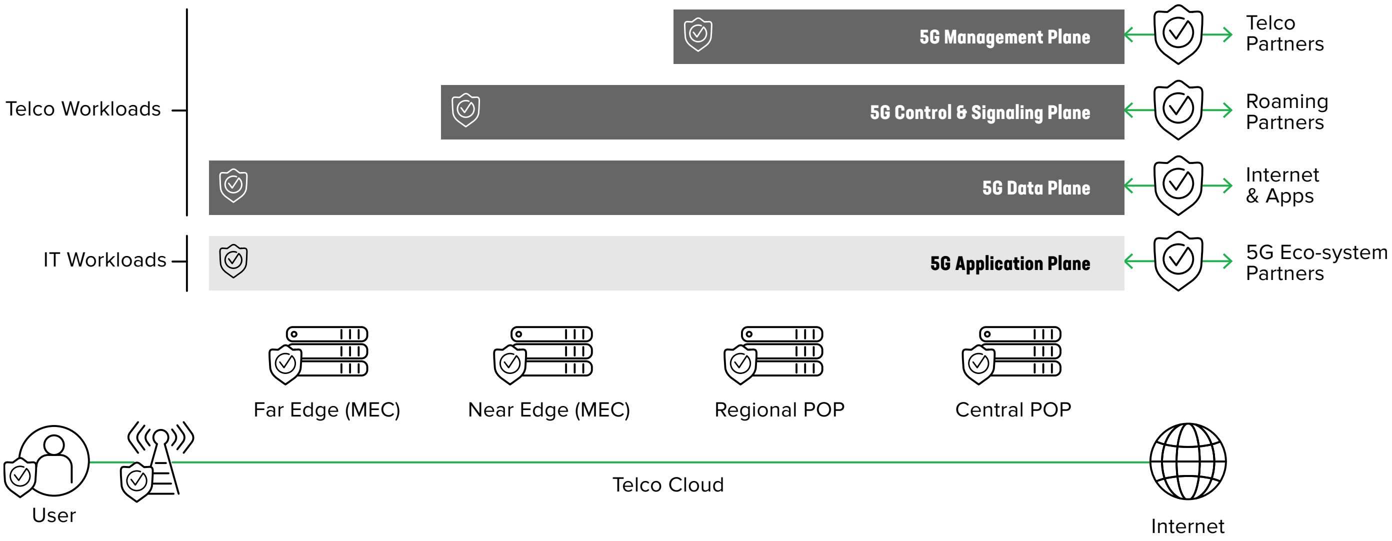


Figure 4: The four planes in the telco cloud that need to be secured

The sections that follow outline some of the key architectural and related security considerations in each plane.

The Data Plane–Seeking Agility and Efficiency

In a 5G network, the data plane is set to carry both a greater volume and a greater diversity of traffic than in 4G networks. In this demanding environment, the traffic management tools need to be both highly efficient and scalable.

SECURING THE N6 INTERFACE

The traffic flowing through a traditional mobile network is managed by the S/GI-LAN interface. In a 5G network, this role is performed by the N6 Interface, which sits between the UPF (user plane function) and the Internet.

This interface provides a number of security services, such as the N6 firewall, carrier-grade network address translation (CGNAT), protection against DDoS attacks, DNS security and, increasingly, dedicated IoT firewalls. In the latter case, a firewall is used to provide very granular level control as to which IoT device can communicate with which server in the back end. That ensures the traffic doesn't leak between different IoT use cases and eliminates leakage to the public Internet.

To give themselves greater scalability and flexibility, many mobile operators are now using a virtual architecture, rather than a physical architecture, to provide these services. They hope that a virtual architecture will make it easier to deploy new network functions and launch new services without any network downtime, boosting service agility. If each of these functions is provided by a different vendor, that agility can be hard to achieve. Relying on multiple vendors means the delivery of new services to subscribers becomes more complex and can involve major delays.

In contrast, employing a single vendor to cover a wide range of different services simplifies the integration into higher-level automation or distribution tools (see Figure 5). Deploying a unified platform also reduces the initial capital outlay because fewer CPUs are required to run a suite of software than an array of separate network functions. A study by one of F5's MNO customers concluded that using a single VNF (virtual network function) instance for S/Gi-LAN services results in a 60% reduction in the total cost of ownership of the necessary computing infrastructure.

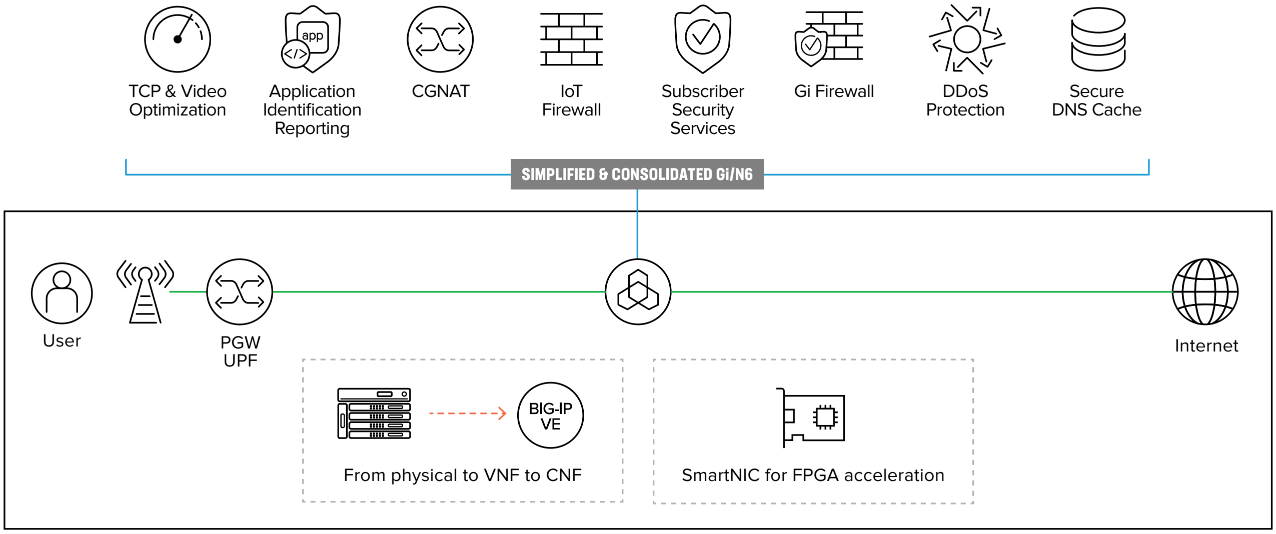

TCP & Video Optimization
app
Application Identification Reporting
CGNAT
IoT Firewall
Subscriber Security Services
Gi Firewall
DDoS Protection
Secure DNS Cache
SIMPLIFIED & CONSOLIDATED Gi/N6
User
PGW UPF
Internet
BIG-IP VE
From physical to VNF to CNF
SmartNIC for FPGA acceleration

Figure 5: How a unified platform can help secure the data plane

DEALING WITH DDOS ATTACKS

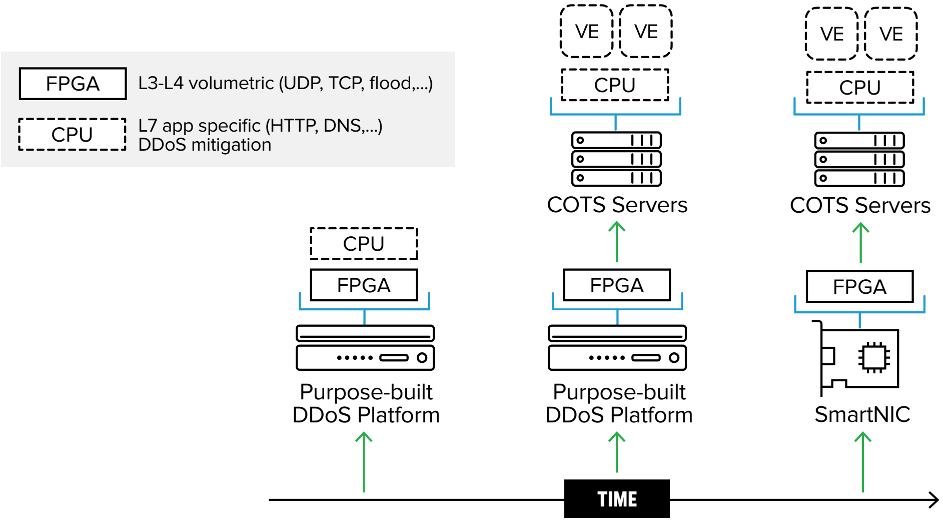
One of the biggest security challenges in the data plane is presented by DDoS attacks. Historically, operators have deployed a purpose-built platform for DDoS mitigation with two major components—a FPGA (field-programmable gate array), which provides hardware acceleration for some functions, and CPUs for very granular processing. The FPGA is programmed to process network traffic and mitigate large volumetric attacks, which use a barrage of malformed packets to consume network bandwidth and large numbers of CPU cycles to bring the CPUs down.

Because operators have moved to virtualized architectures with cloud-based applications, they can deploy X86 server platforms, protected by purpose-built DDoS mitigation platforms with FPGA-based hardware acceleration techniques to deal with volumetric traffic attacks.

This approach takes up valuable space in data centers, which can be an issue in network edge locations where space is in short supply.

These dedicated platforms are no longer necessary with the arrival of the SmartNIC from Intel. A SmartNIC, installed on a standard x86 server, enables F5 to leverage Intel's FPGAs, providing hardware-accelerated DDoS mitigation capabilities in a fully virtual environment (see Figure 6).

Figure 6: Deploying SmartNICs to help protect CPUs from DDoS attacks



FPGA
L3-L4 volumetric (UDP, TCP, flood,...)
CPU
L7 app specific (HTTP, DNS,...) DDoS mitigation
CPU
FPGA
Purpose-built DDoS Platform
VE
VE
CPU
COTS Servers
FPGA
Purpose-built DDoS Platform
VE
VE
CPU
COTS Servers
FPGA
SmartNIC
TIME

This new architecture helps prevent edge computing CPU resources from being overwhelmed by DDoS attacks that seek to consume network bandwidth and disrupt low latency applications, such as video surveillance, machine vision, and remote control systems. Always on, inline DDoS protection offered by this new architecture is critical to protect real-time, latency sensitive applications.

The Control Plane—Adapting a Microservices Architecture

F5'S APPLICATION SECURITY CAPABILITIES, FOR EXAMPLE WEB APPLICATION FIREWALL AND DDOS POLICIES, CAN BE INTEGRATED INTO THE KUBERNETES INGRESS FRAMEWORK.

A modern application is typically built out of different microservices, each handling a specific function of the application, such as order management, reporting, or payments. The application can be packaged in containers, allowing for automated deployment, scaling, and management using the Kubernetes open-source system. A Kubernetes ingress controller is generally used to enable end users to interact with the application. The ingress controller also provides traffic management, ensuring that downstream user requests get routed to the right microservice within the container cluster. F5's application security capabilities, for example web application firewall and DDoS policies, can be integrated into this Kubernetes ingress framework.

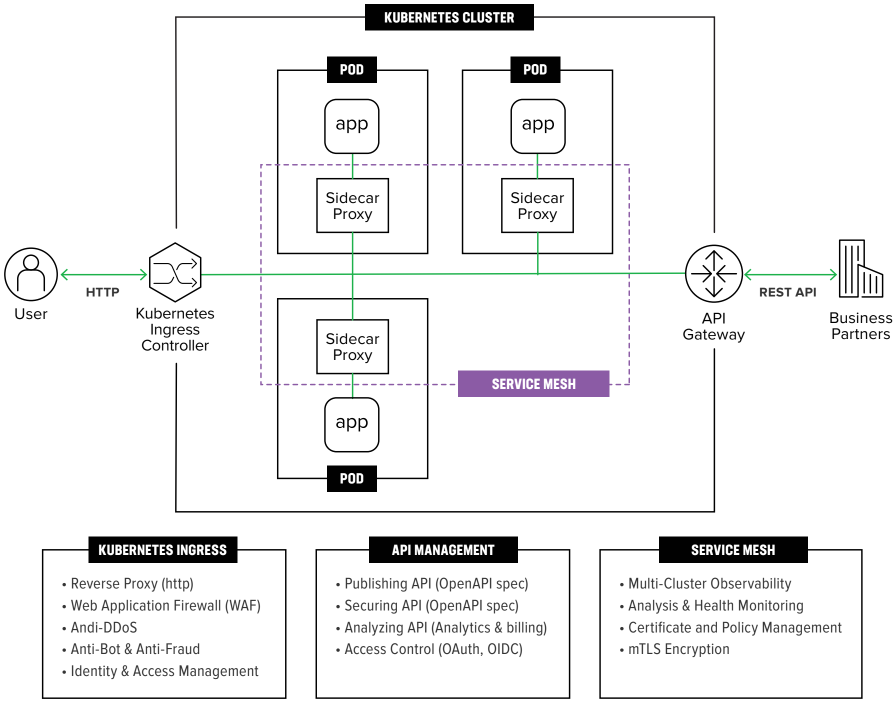

KUBERNETES CLUSTER
POD
app
Sidecar Proxy
POD
app
Sidecar Proxy
User
HTTP
Kubernetes Ingress Controller
Sidecar Proxy
SERVICE MESH
app
POD
API Gateway
REST API
Business Partners
KUBERNETES INGRESS
• Reverse Proxy (http)
• Web Application Firewall (WAF)
• Andi-DDoS
• Anti-Bot & Anti-Fraud
• Identity & Access Management
API MANAGEMENT
• Publishing API (OpenAPI spec)
• Securing API (OpenAPI spec)
• Analyzing API (Analytics & billing)
• Access Control (OAuth, OIDC)
SERVICE MESH
• Multi-Cluster Observability
• Analysis & Health Monitoring
• Certificate and Policy Management
• mTLS Encryption

Figure 7: Modern apps use a microservices architecture and a Kubernetes ingress

Looking upstream, an API gateway can be used to enable an application to securely communicate with business partners. For example, Salesforce.com provides an API that companies can use to extract their sales data and feed it into other business tools, analytics tools, and dashboards. The company providing the API must be able to specify what a business partner is allowed to do with it, as well as implementing authentication and access control.

ADAPTING THE SERVICE MESH

Within the application cluster, the different microservices talk together using a standardized service mesh. This mesh allows the company to observe the traffic flows between these different microservices and, if necessary, protect this traffic using mTLS (mutual transport layer security) encryption.

This IT-based microservices architecture underpins the service-based architecture employed in a 5G core network (see Figure 8).

TELCO SERVICE MESH IS CRITICAL FOR:

- CNF traffic capture at sidecar proxy
- Pre-encryption tapping
- Leverage existing packet broker infrastructure
- Reduce SSL load on brokers

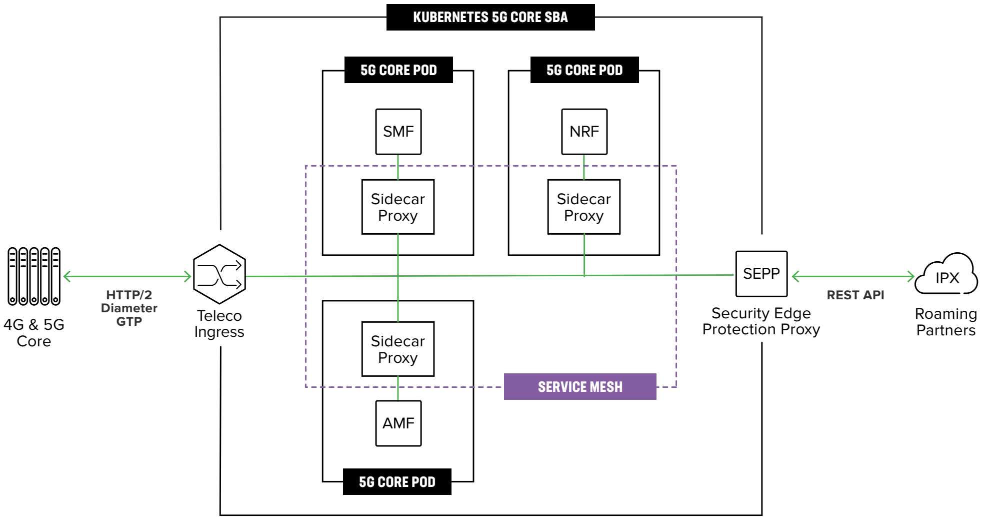


Figure 8: The service-based architecture employed in the 5G core

But there are a few differences, as some specific functionality is required to make these solutions work well in a telecom environment:

The telco ingress requires:

- HTTP/2, PFCP, Diameter and GTP security
- 4G/5G interworking function

The SEPP requires:

- JSON object signing
- Message transformation and traffic mediation

The telco service mesh requires:

- Packet capture (lawful intercept)
- Traffic routing and overload protection

In the case of a 5G service-based architecture, the Kubernetes ingress doesn't interface directly with a user, connecting instead to other 4G and 5G core elements using traditional telco protocols, such as Diameter[1] and GTP[2] where 4G-to-5G interworking is required. In some cases, HTTP/2 messages will need to be translated into Diameter messages.

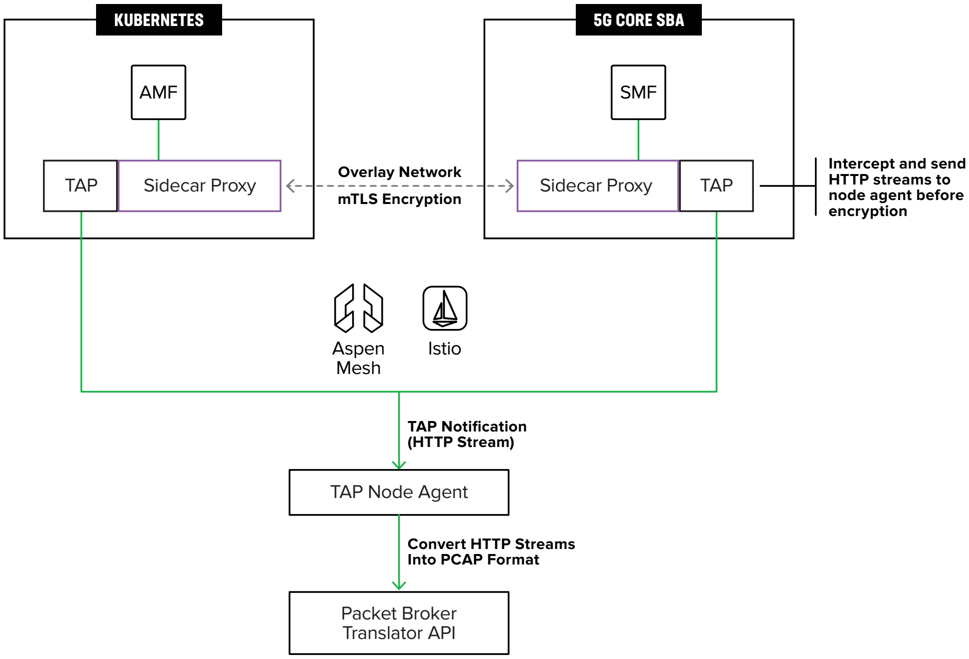
As applications in the control and signaling plane can be exposed to the outside world, security measures will be required. A GTP tunnel, for example, is used to connect the core network of one operator to the core network of another operator to enable end users to roam from one network to another. A malicious attack on the roaming partner's network could compromise the host network.

Defined by 3GPP, the API management interface towards roaming partners is called the security edge protection proxy (SEPP). It is based on the same concepts as those used in IP-centric API gateways, with some additional elements, such as JSON object signing. The SEPP also has some message transformation and traffic mediation capabilities.

In a 5G service-based architecture, a telco service mesh provides interconnectivity between the different 5G core functions. In response to requests from customers, F5 has added some additional management mechanisms, such as packet capture for lawful intercept purposes and traffic routing and overload protection.

In the telco service mesh, you could have a core access and mobility management function (AMF) and a session management function (SMF) talking to each other, for example. These interactions can be secured using F5's Carrier-Grade Aspen Mesh solution, which is based on Istio open source technology (see Figure 9).

Figure 9: Adapting an IT service mesh to meet telco requirements


KUBERNETES
5G CORE SBA
AMF
SMF
TAP
Sidecar Proxy
Overlay Network
mTLS Encryption
Sidecar Proxy
TAP
Intercept and send HTTP streams to node agent before encryption
Aspen Mesh
Istio
TAP Notification (HTTP Stream)
TAP Node Agent
Convert HTTP Streams Into PCAP Format
Packet Broker Translator API

The Management Plane—New APIs Bring New Risks

The management plane enables MNOs to share network capabilities with third parties using APIs. This is a market that is expanding very rapidly. Mordor Intelligence has forecast it will grow at a compound annual growth rate of 23% during the period of 2020 to 2025, as MNOs seek to monetize data about their networks, customers, and utilization. The TM Forum has

MORDOR INTELLIGENCE HAS FORECAST A COMPOUND ANNUAL GROWTH RATE OF 23% FROM 2020 TO 2025 IN THE THIRD PARTY NETWORK SHARING MARKET.

launched initiatives to create open API specifications at the OSS/BSS level, allowing MNOs' business partners to tap into these OSS systems and create their own services based on the information that MNOs have available.

This opportunity is expanding with 5G, which supports a network exposure function, enabling third parties to assume more control over some network expressions. One notable example is slicing: the MNO separates out a portion of its network and dedicates it to a specific use case. The customer can access this slice via an API. While this capability enables new business, these APIs are also opening up new security risks, just as roaming opens up new risks related to roaming partners. Therefore, securing these APIs is extremely important.

The Application Infrastructure–Portable Security Required

The applications residing inside the telco cloud have a number of potential vulnerabilities. They can be subject to attacks on the DNS infrastructure, the SSL services and application protocols, as well as attacks on end-user's clients. Web applications are vulnerable to automated attacks, requiring anti-bot detection and protection.

To effectively protect applications and the users accessing them, identity and access management and an anti-bot enabled web application firewall are required as a minimum. For microservices applications, API security is needed at the back end to eliminate risk. See Figure 10 for overview of the necessary security components.

As telco architectures change with the advent of 5G, it is important that these application security measures are multi-cloud and multi-platform capable. Applications residing in a private data center might move to a public cloud. Later, the MNO might bring them back to their own data center or move them to another cloud.

Therefore, relying on the specific security capabilities provided by these different cloud platforms is not advisable. They may be easy to deploy, but the level of protection will be inconsistent and migrating applications to a new environment could be time consuming and problematic.

Installing cloud- and platform-agnostic security measures allows a MNO to move the solution from one environment to the other without sacrificing security functionalities, minimizing disruption to operation, visibility, and compliance.

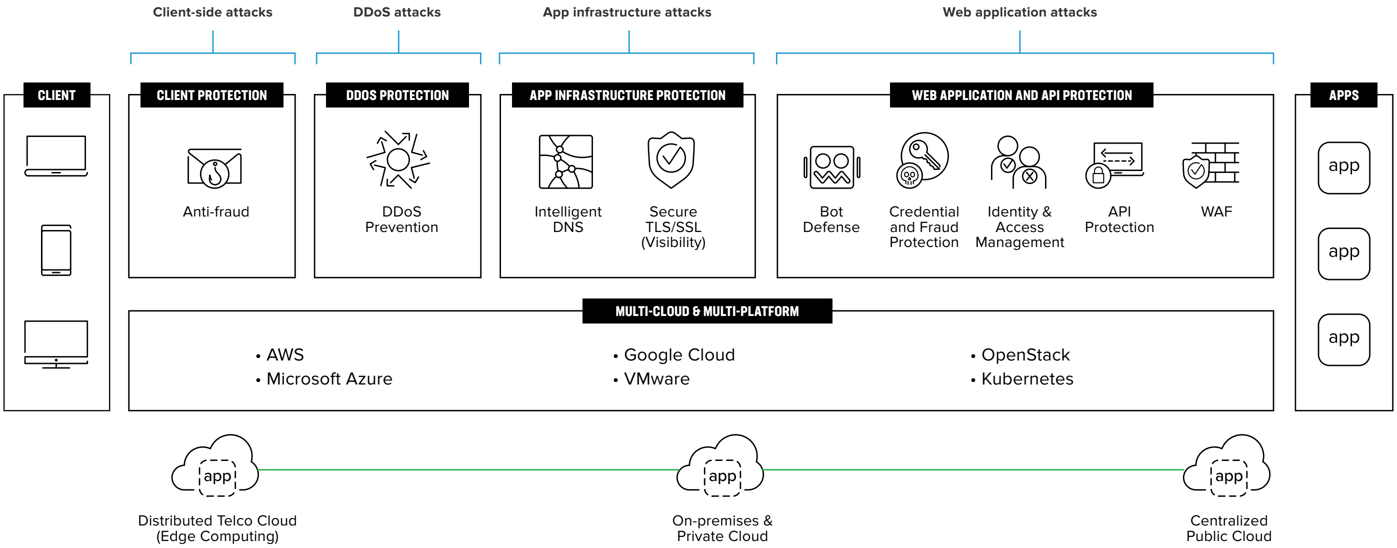


Figure 10: Protecting applications requiires a range of security measures

Conclusion

Although the standards body 3GPP has effectively leveraged IT technology and security mechanisms in the 5G standards, its focus has been on the signaling and control and data planes. In these planes, MNOs can further improve security in a cost-effective way using specialist tools developed at F5.

F5 IS A WELL-ESTABLISHED PROVIDER OF INFRASTRUCTURE, USER, AND APPLICATION SECURITY BOTH FOR ON-PREMISES AND CLOUD DEPLOYMENTS IN THE ENTERPRISE MARKET.

Application infrastructure falls outside the traditional remit of 3GPP. Yet as applications coexist with network functions on the same telco architecture, application security is a key part of 5G security as a whole. As they deploy the telco cloud, operators need to draw on best security practices from the IT sector.

At the same time, MNOs need to be careful that their approach to application security does not mean they are locked into a specific cloud provider. The public cloud market in the telecoms sector is very fluid; MNOs are forging new alliances to deploy the edge computing infrastructure required for 5G.

F5 is a well-established provider of infrastructure, user, and application security both for on-premises and cloud deployments in the enterprise market. F5 has also been a core provider of infrastructure solutions to MNOs for many years. With the advent of 5G, F5 can help MNOs adapt to the new paradigm and deliver 5G-based services reliably and securely.

For more information about 5G security solutions, contact F5.

Endnotes

[1] The diameter protocol provides authentication, authorization, and accounting messaging services for network access and data mobility applications.

[2] The GTP protocol is used to transmit user and control traffic on 2G, 3G, and 4G networks