White Paper

Scaling Mobile Network Security for LTE: A Multi-Layer Approach

Prepared by

Patrick Donegan
Senior Analyst, *Heavy Reading*
www.heavyreading.com

on behalf of

www.f5.com

April 2014

New Traffic Volumes, New Security Threats

With the extensive rollout of 3G High Speed Packet Access (HSPA) and Long Term Evolution (LTE), the mobile network is on course to evolve from its origins as a voice-oriented time-division multiplexing (TDM) network to a full-fledged Internet service provider (ISP) network that provides mobility and voice services almost as a value-added service.

It's not news that many, perhaps most, mobile operators are only barely coming to grips with the scale of network traffic they will need to support as they become ISPs. In developing markets, operators that have only recently started rolling out 3G to the mass market are still seeing the early adopter's data growth rates of 500 percent or 600 percent a year. LTE operators in more mature markets are still often seeing data traffic doubling year on year.

Despite the important efforts of many players in the mobile industry ecosystem, the growth in signaling traffic in the network continues to materially outpace even the growth of user traffic. The mobile network these days is a very heterogeneous, complex environment. Operators are seeing several hundred different applications running in the network, all connected via data flows, all of them interfering with signaling. That doesn't just impact network resources and put service assurance at risk; it also causes smartphone battery drain, again jeopardizing the user experience and even the operator's ability to bill for services.

This tendency for signaling traffic growth rates to outstrip growth rates in data traffic can be expected to continue in the coming years. There has certainly been some important progress in reducing some of the excess signaling generated by some smartphone apps transmitting so-called "keep alive" messages. But the growth of signaling nonetheless continues to outpace user traffic.

Moreover, while smartphone signaling has gained some notoriety in the mobile industry, this is by no means the only source driving up signaling traffic. Over-the-top (OTT) Internet players such as Google, Microsoft and Yahoo! are also adding tremendously to signaling volumes as they deliver more and more multimedia services, including video and gaming apps. As discussed below, some of the changes in the network architecture with the rollout of LTE are also contributing to this problem.

New Security Threats to the Mobile Network

Consistent with the transition they are undergoing, mobile operators have started to be subjected to the kinds of security incidents and attacks that go with becoming an ISP. Some of these can degrade the performance of the network or an individual service. Some can cause outright outages of a network or service. Others can steal highly sensitive and valuable information – either from customers or from the operators themselves.

Some of these security events are benign "attacks" on network resources from any one of a number of signaling overloads of the kind referred to above. These tend to arise from weak security testing; inadequate operational procedures; poor design of the network equipment or network architecture; or misconfiguration of network infrastructure. Several leading LTE operators have suffered high-profile outages from such benign security events in the last couple of years.

These issues arise in the context of the ferocious pace of the 3G and 4G market that can sometimes drive operators to push just that bit too hard against the limits of new networking technology. This intense competition for market share can sometimes leave operators open to vulnerabilities that even some comprehensive lab testing wouldn't uncover before commercial deployment.

Other types of security events result from malicious security attacks carried out by attackers from outside the network or by rogue insiders in the form of current or former employees of the operator or partner companies. For example, *Heavy Reading*'s October 2013 global survey of mobile operators yielded evidence that around 60 percent of mobile operators had experienced malicious attacks that caused an outage or degradation in a major part of their network lasting at least one hour in the previous 12 months. The survey also showed that 34 percent of mobile operators had suffered more than two such incidents. And to make matters worse, it can sometimes be very difficult for operators to distinguish benign traffic anomalies from malicious attacks, thereby rendering the threat landscape all the more challenging to understand, manage and protect against.

The Network Security Architecture Has Been Built Out *Ad Hoc*

In order to protect their networks against the unpredictable and potentially hazardous IP networking environment that is accompanying their transition to being ISPs, many mobile operators have already had to make substantial new investments in the capacity and features supported in their core infrastructure and, specifically, in their network security infrastructure. But this security architecture and infrastructure has tended to be built out *ad hoc*, typically deploying multiple single-purpose hardware platforms from various vendors. Changes have tended to be implemented reactively and tactically rather than with a longer term strategic security framework in mind.

For example, with the launch of 3G a mobile operator might have started out with a basic firewall and network address translation (NAT) gateway to protect the Gi interface at Layers 1-4. Over time, the operator might then have added some deep packet inspection (DPI), some load balancers, perhaps an intrusion prevention system and possibly another firewall at the Gp roaming interface. More recently, the operator may have refreshed its original firewalls with next-generation firewalls with some application layer intelligence, or perhaps some specialized application layer security devices enabling, say, URL filtering.

This *ad hoc* accumulation of security products risks becoming a bottleneck on the scalability of the network, as well as on the overall effectiveness of the security architecture. This white paper provides a blueprint for how mobile operators can stay ahead of the curve of both benign and malign security threats. It examines how operators can adapt and scale their security architecture for tremendous traffic growth on both the user and signaling planes, and it explores how they can dynamically adjust their security posture in line with changes in the security threat landscape, new capabilities in telco networking and changes in customer demand.

LTE Drives Traffic… & Security Vulnerabilities

Operators that are providing 2G and 3G services may already be feeling the strain of the so-called mobile data "tsunami," but as the saying goes, "they ain't seen nothing yet." LTE operators report data consumption anywhere from two to three times as high as 3G data consumption. With that comes the LTE roadmap of ever more powerful smartphones, an ever-increasing proliferation of apps (and a subset of poorly written apps) and an evolution to the so-called Internet of Things, in which billions of devices other than phones and smartphones will be connected to networks.

As depicted in **Figure 1**, LTE materially shifts the goal posts once again in regards to network traffic, security vulnerabilities and the new challenges that face the operator's security architecture.

Figure 1: LTE Drives Faster Traffic Growth as Well as New Security Vulnerabilities

LTE NETWORK CHARACTERISTIC	NETWORK TRAFFIC IMPACT	NETWORK SECURITY IMPACT
Higher bandwidth, faster throughput	Increased data traffic Increased consumption of all network resources	This is a better environment than 3G for attackers launching messaging spam and volumetric DDoS attacks
Flat IP architecture, no RNC any more	All signaling traffic hits the EPC	S1 and X2 no longer natively encrypted RNC no longer there to act as security buffer for the EPC
Greater IMS adoption, including with VoLTE & RCS	Lots more SIP & Diameter signaling traffic	SIP is very easy to use for malicious attacks Diameter has known vulnerabilities in network overload scenarios
Increased use of 3GPP policy tools	Lots more diameter signaling traffic	Diameter vulnerabilities
Small cell proliferation	Increased data traffic	Small cells are vulnerable to physical tampering by attackers

Source: Heavy Reading

Some specific points to note:

- **The higher bandwidths provided by LTE are great for consumers, but they are great for attackers, too.** The more bandwidth at an attacker's disposal, the greater the impact they can have, for example, unleashing a distributed denial-of-service (DDoS) attack or generating messaging spam. The threat potential of the LTE Uplink is relevant here, with speeds of 5-10 Mbit/s now increasingly available to benign user and malicious attacker alike, and the performance set to accelerate still further, including with LTE Advanced (LTE-A).
- **There is no longer a radio network controller (RNC) in the LTE network to serve as a security buffer for the Evolved Packet Core (EPC)**. All the network signaling from a mobile session now hits the EPC, whereas in 3G some of that signaling impact was absorbed by the RNC. An attacker that

penetrates the edge of the mobile network could now have a straight shot at the EPC, whereas the RNC stood in his way previously. And to cap it all, the S1 and X2 interfaces between the eNode B and the EPC are no longer natively encrypted as the equivalent Iub and Iur interfaces were previously in 3G. Outdoor small cells have not taken off in 3G, but they will with LTE, and these wall- or pole-mounted devices are uniquely vulnerable to physical tampering.

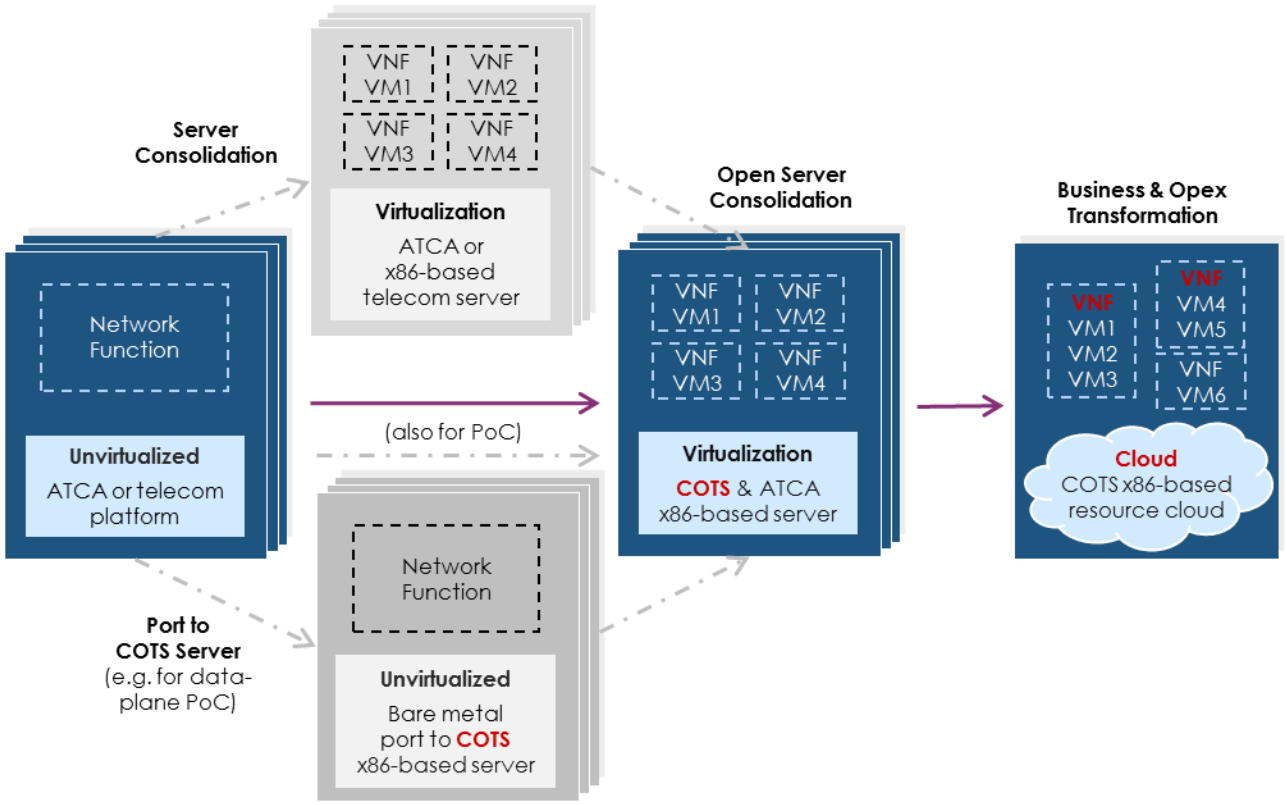
- **The LTE and IP Multimedia Subsystem (IMS) environments – including upcoming VoLTE and Rich Communications Suite (RCS) apps – drive substantial new signaling traffic volumes and expose new security vulnerabilities.** This is because they make extensive use of the Session Initiation Protocol (SIP) and Diameter protocols, both of which are very extensible and IP-based protocols and, hence, a lot more vulnerable than the traditional telecom-oriented SS7 protocols that predominate in the 3G network. The moment SS7 protocols are IP-based and running over Sigtran, they are also much more vulnerable. Although there is ongoing work to fix it in standards bodies, Diameter is still being optimized to handle complex overload scenarios in telecom networks that have already triggered outages in some of the world's leading LTE networks. And because of its high extensibility and openness as an IP protocol, SIP is a preferred protocol for attackers looking to launch attacks, such as DDoS attacks on IP networks. SIP is also considered among the easiest protocols for attackers to use.

Security Is a Lead Use Case for Virtualization

There are many capabilities that are already available to operators to help them evolve their networks for these challenges. Consider the value proposition of network functions virtualization (NFV), which together with software-defined networking (SDN) is one of the two most prominent telco networking concepts of 2013 and 2014.

Virtual editions of a lot of networking software, including security software, are already widely available for use throughout the network. And today's telco network infrastructure often comprises assorted proprietary hardware, which is expensive, challenging to interoperate and inflexible. As shown in **Figure 2**, this paradigm can be complemented and ultimately superseded by network application instances that are virtualized in software and run on any third-party hardware. There are clear examples where this can dramatically improve scalability, as well as dramatically reduce costs.

Figure 2: *Heavy Reading*'s Virtualization Roadmap



Source: Heavy Reading

Virtualization began by being implemented in the enterprise data center and has started to be deployed in telco networks. As it is starting to be implemented in the telco environment, network services that ensure the stability and security of the network are among those that operators, including mobile operators, are most interested in virtualizing. And there are some very good reasons for this:

- **Some traffic-intense interfaces, such as the Gi interface, and some lower-layer line rate DoS protection, will likely remain suited to dedicated security hardware for a long time to come.** Besides these cases, however, if there were ever a telco networking domain or a set of networking capabilities that is ripe for virtualization from a cost and scalability perspective, it is the network security domain. As outlined above, the security environment is among the most dependent on different proprietary hardware platforms that are difficult to interoperate and scale dynamically at low cost. The hardware dependency is even more pronounced in the case of some operators that still rely on enterprise security vendors, some of which renew their proprietary hardware every three years.
- **As is well understood, with the growing involvement of organized crime, nation states and terrorist organizations, the Internet threat landscape is posing ever-increasing risks to critical infrastructure, including telecom networks.** And as has already been shown, far from being immune to these trends, the LTE network roadmap is dismantling many of the remaining telecom-era security defenses that mobile operators enjoyed in the 2G and 3G eras. As a result, the mobile network is becoming a lot more vulnerable to both benign and malign security attacks than it ever was in the past.
- **Some network security threats like many DDoS attacks are associated with a high level of burstiness.** This renders the dynamic scalability associated with virtualization particularly well suited to virtualizing security instances that are designed to thwart them. For example, virtual editions of security software can be spun up *ad hoc* in direct response to volumetric attacks that may only last an hour or two. They can then be torn down again as an alternative to putting a lot of investment in another dedicated hardware platform that will likely be heavily under-utilized. This model of virtualizing security is already seeing significant adoption in the enterprise environment. And leading security vendors are navigating the challenge of re-aligning their software pricing to allow for this kind of "pay by the hour" and other innovate pricing models.

For these reasons – as well as the more high-profile instances of mobile operators virtualizing their mobile packet core – *Heavy Reading* has encountered mobile operators developing use cases for virtualization of their load balancing, domain name system (DNS), DPI, NAT and firewall capabilities. Firewall virtualization in particular appears to be popular among these use cases.

Virtualization therefore presents an important part of the operator's roadmap for scaling and hardening mobile network security in the LTE era, albeit with a couple of very important qualifiers.

- **Virtualization need not be – indeed should not be – any kind of a panacea for scaling network security.** By no means does every security instance need to be virtualized. There are some traffic-intensive interfaces such as the Gi or SGi that will always require large and increasing amounts of capacity dedicated permanently to protecting that interface. And in these cases a large rack of dedicated, proprietary hardware may remain the optimal model for many years to come, particularly through the end of the current generation of the equipment's lifecycle. Rather, virtualization provides the operator with a more flexible set of options as to how, when and where the security architecture is dynamically adjusted to meet the new challenges of the LTE era.

- **It's inherent in the security function that security professionals within the operator are rightly vigilant with respect to the security vulnerabilities associated with altering the networking paradigm in favor of virtualization.** Great care must be taken to set appropriate security rules for the way virtual editions are spun up. Security itself cannot be compromised in the pursuit of greater flexibility and scalability in the security architecture. Important standards work in this area is also being done by the ETSI Special Experts Group, which was established in early 2013 to address security issues with NFV. This group is providing security guidance to other ETSI Working Groups on NFV via a new security reference framework, identifying generic and new security threats that are unique to the networking aspects of virtualization and are advancing the remediation of those threats to the network environment through standardization activity either within other industry standards bodies or within ETSI itself.

The Case for a Multi-Layer Security Strategy

Although the value proposition will certainly be enhanced with ETSI-standard NFV, virtualization solutions are at least available today to support mobile operators scale their network security strategy. Other important pieces of the puzzle are in the pipeline: One high-level idea that can serve as a useful umbrella term for the direction network security strategy needs to take is that of multi-layer security. This assumes a more holistic approach to the security of the mobile network. This embraces every layer and at all points in the network: across the user plane and signaling plane; across the end device, cloud and network; and embracing Layers 1-4 and Layers 5-7.

As discussed above, security policy in the mobile network has evolved in a piecemeal, *ad hoc*, way that has made it very challenging to take this kind of holistic or multi-layer view. Much of the security the operators required for 2G was already baked into 3GPP standards so they simply added in new security capabilities in response to the realization that a specific threat needed to be addressed. Hence the range of network security capabilities itemized above has been built up incrementally and deployed selectively – usually deploying multiple different vendors' products on multiple different hardware platforms.

Most recently, some operators may have deployed specialized DDoS protection in the network, and perhaps some AV clients onto smartphones. In some cases, DDoS protection may have been deployed using one vendor's platform to protect Layers 1-4. A few are deploying protection at Layers 5-7, sometimes using a different vendor. And for LTE, operators are deliberating whether to encrypt the S1 and X2 interfaces using IPsec per 3GPP recommendations – and if so whether to encrypt just the user plane traffic, just the control plane traffic, or perhaps both. They also may be considering how many Diameter routing agents to put in the network and where. And in light of the increased demands being placed on it by both network infrastructure and subscriber application look-ups operators may be reviewing the security and resiliency of other key infrastructure, such as the DNS.

This *ad hoc* accumulation of network security capabilities, leveraging multiple different proprietary platforms, has consequences. And one of these consequences is that a common pain point in the mobile network today is that the network and security operations teams are overwhelmed by huge volumes of single-purpose, one-dimensional data feeds that trigger alarms in isolation from other events elsewhere in the network. As signaling volumes grow, so the volume of signaling events triggering alarms is also liable to grow.

Today mobile operators lack a means of correlating all the many different security and security-impacting data feeds, alarm mechanisms and mitigation solutions looking for different threats in so many different points in the network. And while large operators may be well-resourced enough to determine their emerging security requirements, as they scale up the majority of small- and medium-sized operators are not.

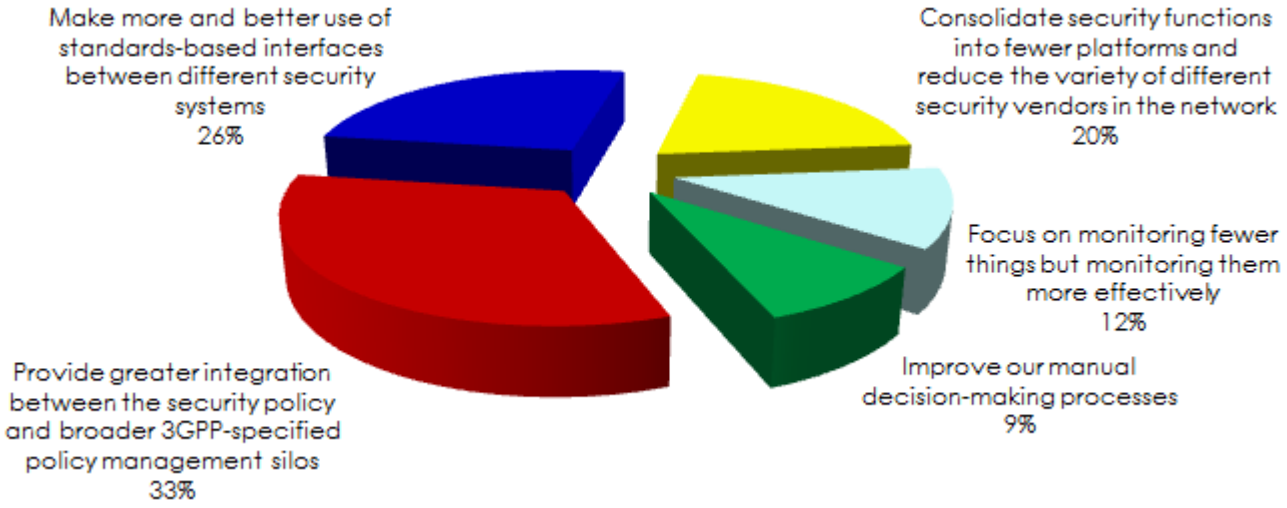
A multi-layer security strategy with a holistic end-to-end view of the network's security status would not only support dynamic scaling of the security infrastructure so that new threats can be anticipated and defended against in minutes or hours rather than in days, weeks or months, but it would also enable mobile operators to correlate security policy with subscriber policy in real time to better tailor network resources to the unique requirements of each individual end user and each and every session.

Leveraging 3GPP's Policy Environment for Dynamic Security Policy

As shown in **Figure 3**, *Heavy Reading* research demonstrates that mobile operators recognize the need for a more unified view of security issues in the network and associated mitigation options. Consistent with virtualization trends, many believe that consolidating security functions into fewer different platforms would be helpful toward that end. Many also believe in making better use of standards-based interfaces between different security elements. But as shown in **Figure 3**, what operators clearly believe can most help them achieve a more unified view for addressing security threats is greater integration between the security policy and the 3GPP policy management environment, which is driven by the Policy Control Resource Function (PCRF) and Policy Enforcement Points (PEP).

Figure 3: Priorities for Creating a More Unified View of Security Threats

What are your company's priority strategies for better correlating the multiple feeds of security-related data that you receive into a more unified view of security threats and associated mitigation options?



Source: Heavy Reading's Mobile Network Security Survey, October 2012; n=83

As LTE is rolled out, one barrier to tailoring security policy to the evolving threat environment – one might even think of this as a vulnerability in its own right – is that security policy is by its nature static. Security rules, for example, on firewalls, are generated and applied in the operations environment until such time as they outlive their usefulness and new rules are generated and applied. But if the security policy environment can be subscriber- and application-aware through interaction with the 3GPP policy environment, the opportunity opens up for dynamic security policy in which policy changes dynamically in real time, per user or per application, based on changing network conditions or requests from users.

What this means, in terms of use cases, is that different security policies such as intrusion detection, URL filtering or malware detection can be dynamically applied per application or subscriber. So instead of these security policies being universally applied to each and every application and subscriber, they can be selectively applied to specific subscribers or applications, drawing on real-time intelligence on the application, prevailing network conditions and subscriber preferences.

Subscriber awareness would also extend to allowing the operator to vary security policy dynamically according to which device the subscriber is using – for example, a higher level of threat detection might be warranted in the case of Android devices rather than iPhones. Or at an even more granular level, different security policies might be applicable according to which specific Android device is being used. The only way for security policy to evolve in this direction is through greater subscriber and application awareness, and the obvious source of that information is the 3GPP policy environment.

There are certainly some practical barriers to dynamic security policy, and organizational challenges are foremost among those. In the first place, the 3GPP policy architecture has not been designed to align with or support network security policy so for it to fulfill this role would likely require changes in the 3GPP standards.

Consistent with current 3GPP standards, core networking and security teams typically also have separate remits and are held accountable for them. While some form of convergence between the two environments can support one another's respective goals, organizational boundaries aren't always so easily broken down.

Dynamic security policy also requires significant involvement – in some cases leadership – in security policy by the mobile operator's sales and marketing organizations, something that is pretty much unprecedented. Aligning the perspectives of these three very different organizations within the operator will likely prove challenging in many environments. The benefits must be equally well understood by all parties, and any risks are mitigated.

Service Chaining With NFV

Service chaining is the oldest telecom network concept in the world in the sense of enabling network services like firewall by connecting them between different network elements. But with NFV a new model of service chaining is evolving in which interactions between network elements are abstracted from multiple different proprietary hardware types and managed in software. The idea is that this should allow savings in network resources. It should also materially reduce the risk of configuration error in the network service chain, whether that chain includes firewall, intrusion detection, malware protection or any other network service.

One use case for this kind of service chaining occurs when deploying firewall and load balancing together. With the trend toward virtualization and cloud-based networking, load balancing and cloud security requirements are growing. For example, mobile operators are already virtualizing network nodes such as the EPC, mobility management entity (MME), home location register (HLR), home subscriber server (HSS), authentication, authorization and accounting (AAA) and IMS.

One of the risks that operators have traditionally faced when deploying firewall functionality is that traffic throughput slows as result. This scenario can be avoided when firewall functionality is deployed with load balancing in a service chain. If the solution is designed and implemented in the right way, it's even possible for the operator to achieve gains in throughput with this approach. Another example could be that when decrypting and encrypting IPsec packets on the S1 or X2 interfaces, service chaining could also allow threat detection to take place at the same time according to a service chaining principle. The alternative is to encrypt and decrypt twice for the separate services, which uses more network resources and could also impact latency.

Summary

The way the network security domain has been pieced together over time has served mobile operators well enough until now. By and large, it has enabled operators to "keep the lights on" and minimize the impact of network incidents on subscribers. But the business model and technology environment are changing fundamentally today, and if the network security domain doesn't evolve with that, it will risk compromising operators' ability to protect the network, maintain stability and manage millions of real-time sessions, without costs spiraling out of control.

With LTE, the network must scale to keep pace with traffic growth and, particularly, strong growth in signaling traffic. The threat landscape is evolving to generate increasingly threatening outcomes, be they from malicious attacks or benign network incidents, and the technology and architecture of the mobile network is transforming. The seismic shift to all-IP networking protocols is exposing the operator to more threats, while virtualization provides new opportunities to render the security architecture more flexible and ultimately more robust and effective.

A network security domain that is fit for purpose in the emerging LTE era will increasingly require a multi-layer approach across all the key layers and locations in the network. According to each operator's roadmap and commercial availability of solutions that fit their needs, a dynamic, multi-layer approach will need to leverage virtualization and service chaining, as well as real-time subscriber awareness from the 3GPP policy environment.

About F5

F5 helps organizations and service providers seamlessly scale cloud, core network, and SDN/NFV deployments to successfully deliver services to anyone, anywhere, at any time. F5 solutions broaden the reach of IT through an open, extensible framework and a rich partner ecosystem of leading technology and network orchestration vendors. F5 enables service providers to secure, optimize, and monetize their networks by leveraging contextual subscriber and application information to provide the ultimate customer experience, maximize network efficiency, and deliver services quickly and more cost-effectively. This approach lets customers pursue the infrastructure model that best fits their needs over time. For more information, go to f5.com.