White Paper

Threat Detection & Security Policy Enforcement: A New Inflection Point in Mobile Security

Prepared by

Patrick Donegan
Chief Analyst, Heavy Reading
www.heavyreading.com

on behalf of

www.juniper.net

June 2016

Inflection Points in Mobile Network Security

There have been three major inflection points in the history of security for digital mobile communications services at the network level. The first was the introduction of end-to-end encryption, without which GSM, 3G and 4G would never have grown to be the global, mass-market service that they are today. Less seismic in its impact, but nevertheless significant, was the introduction of firewalls at the key Gi interface with the rollout of GPRS and CDMA 2000.

On a par with second inflection point was the introduction of the new 3GPP Security Gateway (SecGW) with 4G LTE. This gave mobile operators the option to secure S1 interface traffic with IPsec from the eNodeB across the backhaul to the Evolved Packet Core (EPC) in the SecGW. In 2G and 3G, mobile operators didn't need this option because encryption from the handset to the Base Station Controller (BSC) or Radio Network Controller (RNC) was already built into the cellular industry standards.

Toward a Fourth Inflection Point in Mobile Network Security

As shown in **Figure 1**, a couple of key trends in the cyber-threat landscape, in mobile network standards, and in telecom networking more generally, are now driving mobile network security to a fourth major inflection point.

Figure 1: The Four Inflection Points in Mobile Network Security

Year	Landmark	Impact
1991	The A5/1 algorithm for GSM	First ever mass-market communication tool with end-to-end encryption
Late 1990s	Firewall at the Gi Interface	First perimeter security for GPRS & CDMA 2000 cellular data networks
2010	3GPP Security Gateway (SecGW)	A new optional encryption and authentication device between backhaul (S1/X2) and core for 4G
2016	Threat detection and security policy enforcement with SDN & NFV	A more robust, automated, security architecture in which the creation of security policy is more centralized but threat detection and policy enforcement are more distributed

Source: Heavy Reading

- **The threat landscape is becoming increasingly threatening.** This is in terms of the resources going into cybercrime; the sophistication of the attacks in terms of their ability to avoid detection; and their impact in terms of stolen or publicly exposed information and the crippling of network resources.
- **Going back many years, and driven by 3GPP standards, the mobile network architecture has become a flatter, less layered, all-IP network.** One driver for this has been to enable greater distribution of network functions. Consistent with that, "Distributed EPC" was the primary term used during 2008-2013 to describe deploying EPC elements out at aggregation sites in the backhaul to save on backhaul costs and reduce latency.
- **The transformation to software-programmable networks that all communications service providers (CSPs) have begun embarking on recently with**

software-defined networks (SDN) and network functions virtualization (NFV) has now superseded those relatively basic ideas for distributing network features. Hence we no longer think just in terms of distributing monolithic instances of integrated vendor EPC hardware and software. Now we also think in terms of virtual EPC (vEPC) instances that can be spun up anywhere in the network on COTS hardware. In that sense, while SDN and NFV introduce the mobile operator to a host of brand-new options in telecom networking, they also provide better tools for executing on feature distribution objectives that have been in place for a number of years.

Threat Detection & Security Policy Enforcement in the Mobile Network

The latest inflection point in mobile network security consists of large-scale centralization of security policy enabled through more open interfaces driven by SDN. It also consists of greater distribution of threat detection and security policy enforcement throughout the mobile network enabled by both SDN and NFV.

This may not appear as epoch-making as the introduction of encryption for the mass market. However, when you consider the difference that success or failure will make to a mobile operator's competitiveness relative to today's business model as well as emerging opportunities, it is potentially just as profound. And in terms of the design, implementation and operational aspects that the mobile operator has to oversee itself, rather than just fall back on capabilities that are already built into the standard, this fourth inflection point is easily the most challenging to date.

This paper addresses some of the key aspects of this fourth inflection point in mobile security. It describes general changes in the threat landscape, common principles being applied in evolving all ICT infrastructure, and how these apply in the specific context of the mobile network. It makes the case for centralizing security policy while also distributing threat detection and security policy enforcement throughout the mobile infrastructure. It describes the evolution toward that architecture now, while emphasizing the added impetus that the Internet of Things (IoT) and 5G will also bring to executing on these requirements. And it depicts the role of SDN and NFV as enabling technologies in this evolution of the mobile security architecture.

The Old Network Security Model Is Broken

There are two key principles that underpin modern approaches to network security in general, whether that be in the enterprise or service provider environment. The first is that just securing the perimeter against external attacks no longer works.

The sophistication of attacks these days, especially those at the application layer, will inevitably enable some to escape detection by perimeter security devices, no matter how good those perimeter devices are. So in addition to traditional "outside-in" attacks, service provider networks are increasingly vulnerable to "inside-out" attacks in which network elements become infected with malware (BOT malware, for example) and launch attacks on external network elements. They're also increasingly vulnerable to "inside-In" attacks, in which compromised network elements launch attacks on other elements within their own network.

"Inside-out" attacks can damage the end target network as much as the network from which the attack originated, for example by damaging the originator's IP reputation among its peers and risking IP address blocking. This results in blocked access

for users or groups of users where CSPs share a limited public IP address pool across large numbers of users as in a typical mobile network environment.

In the enterprise context, mobile and wireless communications have themselves played a key role in undermining perimeter security. Employees carrying smartphones now flit seamlessly between the "internal" enterprise network and the "external" network environment. Any malware that they happen to be carrying risks infecting either side of what was a much more reliable security divide in the pre-smartphone era. And smartphones or laptops connected via cellular or WiFi require remote access to corporate data when they are "outside" the corporate network, on what was traditionally thought of as the "untrusted" side of the perimeter.

A zero-trust philosophy is the corollary to the principle that perimeter security alone is now ineffective. In today's networking environment, and given the increasing sophistication of the cyber threat landscape, pretty much any network element is increasingly vulnerable to being compromised via multiple different threat vectors delivered from multiple different points in the network.

A robust security architecture traditionally required the ability to detect and mitigate threats at obvious points of vulnerability, such as critical interfaces to the external Internet. Increasingly, a robust architecture now requires the ability to distribute detection and mitigation capabilities beyond these traditional security "hot spots" to the whole network. This is because distributing threat detection and security policy enforcement closer to the source of the threat provides a better means of containing the amount of damage the attack inflicts when it is executed.

A Mobile Security Model That's Breaking Down

In addition to using many open standard IP protocols, the mobile network still uses some unique, internal networking protocols such as GTP and SCTP. The presence of these latter protocols – which are standards-based but nevertheless confined to the cellular environment – does require additional investment on the part of attackers to execute some attacks successfully as compared with the more fully open wireline CSP and enterprise network environment.

Although this does provide an additional barrier to attackers targeting the mobile network as compared to the wired network, it's debatable just how significant that deterrent is today, let alone in the medium term, as attackers' focus on the mobile network inevitably increases and 5G networks are rolled out.

Although mobile operators still need to enforce perimeter security as robustly as possible, the principle of a zero-trust environment is nevertheless just as relevant to the mobile network as any other network environment. As shown in **Figure 2**, in Heavy Reading survey research, mobile operator respondents reported seeing most attacks originating from the Internet and attempting to enter the network via the Gi or SGi interface. But 46 percent of respondents also reported seeing attacks coming from compromised subscriber devices in the RAN.

In the same November 2014 survey, when asked to identify the best place to stop a DDoS attack originating from the subscriber side of the mobile network once it is detected, 56 percent of mobile operator respondents pointed to the RAN or subscriber device, compared with just 26 percent who nominated the 3G mobile packet core or 4G EPC (e.g., via a GTP firewall deployed there) as the optimal place.

Figure 2: Attacks Are Coming From Several Places Now

	The Internet (Gi/SGi)	Compromised Subscriber Devices	Roaming Peers (GRX/IPX)
Most attacks come from here	73%	14%	12%
Some attacks come from here	17%	46%	39%
Few if any attacks come from here	10%	40%	51%
Total	**100%**	**100%**	**100%**

Source: Heavy Reading's Mobile Security Survey, November 2014 #64

The former of these two data points demonstrates that the security holes and vulnerabilities in perimeter security that are evident throughout the ICT infrastructure are also evident in the mobile network: more attacks are originating from "inside" the perimeter. The second data point is quite striking: only 26 percent of respondents consider that the most common approach to protecting against subscriber-side DDoS attacks – protection in the core – is still the best way to deal with these attacks. A clear majority of mobile operator respondents are looking for a more flexible and decentralized security architecture that allows these attacks to be dealt with further out at the edge of the network when malicious behavior is first detected.

Smartphone Botnets Are Getting "Better" All the Time

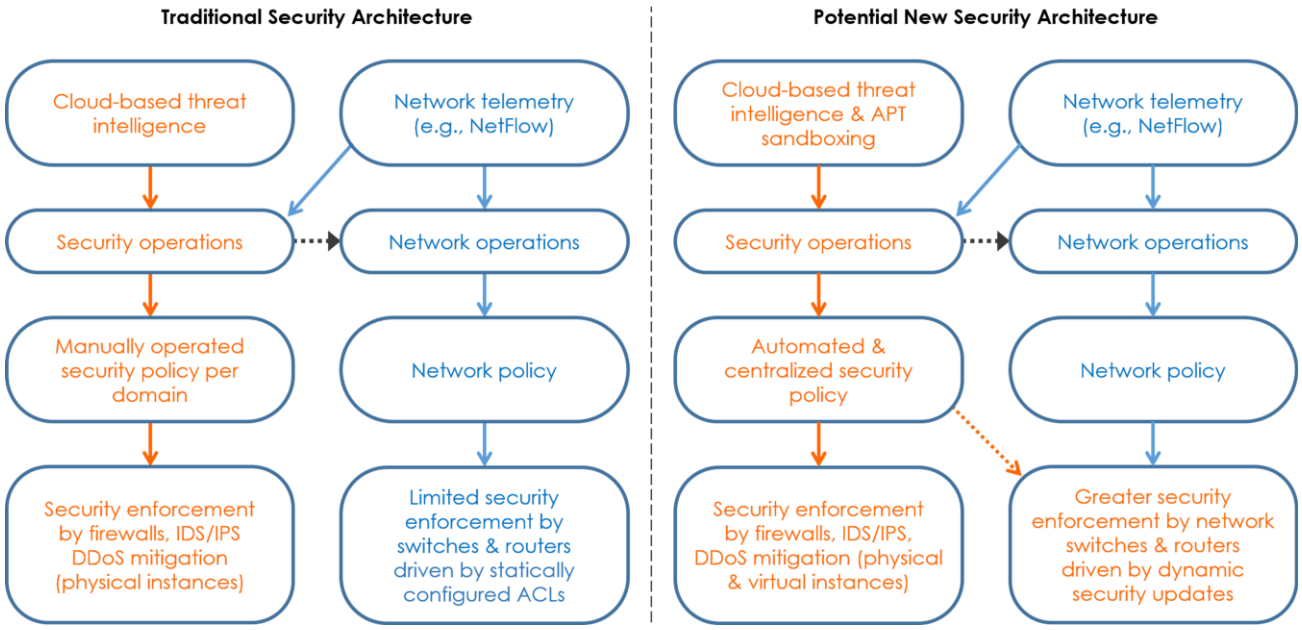
After several years of amateurish efforts compared with the lethal impact of PC and server-based attacks, smartphone botnet malware is now well on the way to presenting a significant threat to mobile network uptime and the privacy of end-user data. The November 2014 discovery of "NotCompatible.C" for Android was a landmark in smartphone botnet software development because it supported sophisticated command and control and encryption to avoid detection – attributes that had traditionally only been seen in Windows-based botnet malware.

A Distributed Architecture for Mobile Security

When referring to security policy in this section we refer to a mobile operator's instantiation and service chaining of specific security applications such as firewall, IDS/IPS or encryption on specific interfaces or traffic streams; the mobile operator's specific rules associated with some of those instances (e.g., firewall rules determining which ports should be open and which closed); rules relating to which network devices may or may not communicate with one another or with the external Internet; and the operator's approach to detecting and mitigating known threat signatures as well as anomalous network or application behavior.

The traditional security architecture on the left of **Figure 3** depicts the way in which network policy and security policy tend to be operated mostly independently of one another in most CSP environments, including in the mobile network. Security policy has tended to draw on threat intelligence in order to inform decisions which are then enforced by the dedicated layer of security infrastructure. As shown, this dedicated layer of security infrastructure includes, but is not limited to, products such as firewalls, IDS/IPS, and DDoS mitigation.

Figure 3: Toward a More Distributed Security Architecture



Source: Heavy Reading

Network Routers Already Help Mitigate Some DDoS Traffic

Even in the traditional security architecture, there has always been a fairly small contribution made by the basic network infrastructure layer to the enforcement of security policy. Most typically, network switches and routers already play a supporting role in DDoS mitigation.

This typically consists of dealing with well-known, low-level threats with easily identifiable signatures, while leaving the more challenging security threats to be dealt with by specialized security gear such as DDoS protection equipment. There is more than an "all hands on deck" justification for offloading a portion of lower-level security functions to the basic infrastructure layer in today's model: Dedicated security products are more expensive than basic routers, so the more threats that can be dealt with by lower-cost infrastructure, the better.

It's notable that in the traditional architecture, the switches and routers that execute on that minor security enforcement role typically don't have direct access to threat intelligence in the form of feeds or dynamic information about potentially infected or dangerous endpoints in the network. Instead, they take part in security policy enforcement with statically configured access control lists (ACLs).

"Dotted-Line" Reporting Between Security & Network Infrastructure

The right-hand side of **Figure 3** shows the case for a new architecture to improve and grow the role of the basic infrastructure in security policy. This can be done in a way that is wholly aligned with the requirements for greater distribution of network security applications. It also aligns with the broader trends favoring feature distribution,

more open network interfaces, and flexible automation of networking applications – both security applications and other network applications – as virtual network functions (VNFs).

There are four fundamental changes in the emerging security architecture as it relates to equipping basic infrastructure elements with the intelligence and flexibility they need to make the contribution that will be required of them in the coming years.

- **The introduction of a direct "dotted line" between the security policy environment and the basic network infrastructure.** This enables routers and switches to be programmed to respond to a great many more threats than they are able to respond to today. It also enables them to enforce a great many more different types of security polices, at many more distributed points in the network, compared with what can generally be achieved with the basic infrastructure today.
- **Virtualization of network security Instances.** In addition to allowing additional security policy features to be enforced by distributed switching and routing elements for the first time, the architecture on the right of **Figure 3** assumes that security VNFs are also being spun up and distributed wherever the presence of threats requires them to be spun up – and on COTS hardware. Most obviously, malware can be detected and mitigated more rapidly, and its impact more effectively contained, if firewall or other security VNFs can be spun up in the closest possible proximity.
- **Automation of security policy.** If the security organization of a mobile operator is to have any chance of getting on the front foot and spend more time anticipating and preparing to combat the upcoming threats that are in the pipeline, then large swaths of the routine administrative work they are currently burdened with needs to be taken off its hands via automation of security policy.
- **The addition of sandboxing to address malware including advanced persistent threats (APTs).** This tends to be less commonly used in the mobile network than in other network security domains today, but we expect the requirement for detecting and safely detonating malware will increase over time as cyber threats, including those focused on the mobile network, increase.

SDN & NFV Are Key Enablers of Mobile Security

SDN and NFV are critical enablers of evolving toward a more flexible security architecture for the mobile network of the kind depicted in **Figure 3**. The open interfaces supported by SDN are key to greater sharing of threat intelligence across network elements and the extension of uniform security policies across core, transport and RAN domains in the mobile network. The lower-cost potential and greater flexibility associated with NFV are critical to enabling distributed enforcement of security policy via security VNFs throughout the mobile network in concert with the security-enabling of the basic network infrastructure.

At the same time that they enable new security threats to be mitigated faster and more efficiently, software programmable networks will undoubtedly introduce some very important new vulnerabilities to the mobile network, as well. For example, the replacement of proprietary with open interfaces with SDN, and the potentially much larger single point of failure arising with both SDN controller and hypervisors. These and other attributes of software programmability threaten a CSP's security stance at the same time as they strengthen it.

As shown in **Figure 4**, the good news is that as that while CSPs clearly recognize that SDN and NFV do represent a threat from a security perspective, they nevertheless see more opportunity than risk. They recognize that SDN and NFV are central to evolving their security architecture with greater flexibility and automation.

Figure 4: SDN & NFV – A Security Threat or an Opportunity for CSPs?

	NFV	SDN
Mostly a threat	8%	15%
Equally a threat & an opportunity	43%	37%
Mostly an opportunity	49%	48%
Total	**100%**	**100%**

Source: Heavy Reading Survey, May 2015 # of CSP respondents: 97

IoT & 5G Use Cases

The proliferation of connected "things" in Internet of Things (IoT) use cases is serving to drive the requirements for distributed enforcement of security policy. Since many end devices in IoT use cases, such as sensors, don't have anything like the compute or power budget required to support endpoint security, network-based security will be key to securing them properly. The closer to those vulnerable end points the mobile operator's detection and security policy enforcement policies can be deployed, the more robust the mobile operator's security stance can be.

5G, which will start to be commercialized within the next three years or so, will also drive security requirements further in this direction. Take network slicing, for example, which is a key element in the 5G value proposition that will differentiate it from 4G and 4.5G capabilities. Network slicing will enable mobile operators to deliver differentiated wireless connectivity services to different customers in specific locations, for varying periods of time, and with unique characteristics such as capacity, speed, robustness, availability and security.

Delivering up the unique security requirements for each network slice in a 5G network environment will be much easier, and will serve up a much better outcome, in a network where the operator is able to draw upon the resources of a distributed architecture for security policy enforcement.

Summary

We are approaching a new inflection point in the way in which security needs to be built into the mobile network. Rather than making do with just a dedicated security overlay infrastructure, security now must be built in throughout the network infrastructure. This requires a centralized security policy layer that leverages cloud-based threat intelligence and is supported by detection and enforcement. This is required in the dedicated physical security infrastructure; in switches and routers; as well as in virtualized instances deployed throughout the network, including in the aggregation and access layer out at the edge of the mobile network.

As security threats increase, as competition between mobile operators shifts from new customer acquisition to customer retention, and as mobile operators target more and more industry verticals with increasingly sophisticated privacy and security requirements tied to their own unique service and application requirements, quality of security is becoming the new QoS. Mobile operators that differentiate here will be best positioned to protect their traditional lines of business and succeed in new vertical markets. Those that don't will see their competitiveness decline.

About Juniper

Juniper Networks (NYSE: JNPR) is in the business of network innovation. From devices to data centers, from consumers to cloud providers, Juniper Networks delivers the software, silicon and systems that transform the experience and economics of networking. The company serves customers and partners worldwide. Additional information can be found at www.juniper.net or connect with Juniper on Twitter and Facebook.