CSR C

Communications Security, Reliability and Interoperability Council

September 2018 WORKING GROUP 3

Network Reliability and Security Risk Reduction

**Final Report – Report on Best Practices and Recommendations to Mitigate Security Risks to Emerging 5G Wireless Networks** **v14.0**

Table of Contents

1 Executive Summary 1-6

2 Introduction 2-1

2.1 CSRIC Structure 2-1

2.2 Working Group 3 Team Members 2-1

3 Introduction to 5G 3-1

3.1 Who Is Defining 5G? 3-1

3.1.1 Status of 5G Standards 3-2

3.2 The 5G Mission 3-3

3.2.1 Speeds and Connectivity in 5G 3-5

3.2.2 Applications Supported by 5G 3-6

3.2.3 The Internet of Things 3-8

3.3 5G Architecture 3-9

3.3.1 5G New Radio (NR) 3-11

3.3.2 Multi-access Edge Computing (MEC) 3-15

3.3.3 Network slicing 3-16

3.3.4 Cloud and virtualization 3-19

4 3GPP SA3 (Security) Standards 4-1

4.1 Increased Home Control 4-1

4.2 Unified Authentication Framework 4-1

4.3 Security Anchor Function (SEAF) and Authentication Framework 4-1

4.4 Serving Network Authentication 4-2

4.5 Mitigation of Bidding Down Attacks 4-2

4.6 Subscriber Identifier Privacy 4-2

4.7 Overview of Security Architecture 4-3

4.7.1 Security Edge Protection Proxy (SEPP) 4-4

4.8 New Security Requirements on the user equipment 4-5

5 The Internet of Things (IoT) in 5G 5-1

5.1 Previous Work Efforts 5-1

5.1.1 CSRIC IV, WG5 Remediation of Server-Based Distributed Denial of Service (DDoS) Attacks 5-1

5.1.2 CSRIC III, WG7 Botnet Remediation and the Anti-Botnet Code of Conduct 5-1

5.1.3 CSRIC II, WG8 ISP Network Protection Practices 5-1

5.1.4 Communications Sector Coordinating Council White Paper 5-2

5.1.5 Report to the President on Enhancing the Resilience of the Internet and Communications Ecosystem Against Botnets and Other Automated, Distributed Threats 5-2

5.2 IoT Service Enablement in 5G 5-2

5.3 Examples of IoT applications 5-3

5.3.1 Industrial Process Automation 5-3

5.3.2 Autonomous Driving 5-3

5.3.3 Smart Grids 5-3

5.3.4 Remote Surgery 5-4

5.3.5 Media and Entertainment 5-4

5.4 Taxonomy of Use Cases 5-4

5.5 GSMA IoT Model 5-6

5.5.1 Service Ecosystem 5-7

5.5.2 Endpoint Ecosystem 5-7

5.6 IoT Connectivity in 5G 5-7

5.6.1 How NB-IoT Flows 5-8

5.7 Risks of IoT in 5G 5-9

5.7.1 IoT Service Assets That Network Operators Can Protect 5-9

5.8 Threat Assessment for IoT 5-12

5.8.1 Emerging Security Threat Environment in the 5G-Based IoT Service Oriented Architecture 5-13

5.8.2 REST API Threat Environment 5-15

5.8.3 Potential Attacks and Security Countermeasures to Protect the 5G IoT Middleware architecture 5-16

5.8.4 API Security Risks in 5G for IoT 5-18

5.8.5 API Security Best Practices to Mitigate Risks to IoT in 5G 5-20

5.9 Improvements to security in 5G for IoT 5-23

6 NFV and SDN 6-1

6.1 Risks of NFV and SDN in 5G 6-7

6.2 Threat assessment of NFV and SDN in 5G 6-9

6.3 Improvements in 5G for Securing NFV and SDN 6-11

7 Open-Source 5G Software Platforms 7-1

7.1 Risks of Open Source in 5G 7-2

7.2 Threat Assessment for Open Source in 5G 7-3

7.3 Threat Mitigation for Open Source in 5G 7-3

7.3.1 Node Security Project (NSP) 7-4

7.3.2 RetireJS 7-4

7.3.3 OSSIndex 7-4

7.3.4 Dependency-Check 7-4

7.3.5 Bundler-Audit 7-4

7.3.6 Hakiri 7-5

7.3.7 Snyk 7-5

7.3.8 Gemnasium 7-5

7.3.9 SRC:CLR 7-5

7.4 Improvements in Open Source as a Result of 5G 7-5

8 Supply Chain in 5G 8-1

8.1 Supply Chain Background 8-1

8.2 5G Supply Chain Ecosystem 8-2

8.3 Supply Chain Critical Infrastructure 8-3

8.4 Supply Chain Vulnerabilities in 5G 8-4

8.4.1 Inadequate Information Sharing across the Supply Chain 8-4

8.4.2 Challenges Specific to Small and Mid-Sized Businesses 8-4

8.4.3 Visibility Risk 8-4

8.5 Threat Mitigation 8-5

8.5.1 Business Management 8-5

8.5.2 5G Supply Chain Threats/Risks: 8-6

8.6 Supply Chain Risk Management (SCRM): 8-9

8.6.1 NIST Information Communication Technology Risk Management Approach: 8-11

8.6.2 4 Pillars of Supply Chain Risk Management: (Derived from ISO SP 800-61) 8-12

8.6.3 Business Management 8-13

8.6.4 Vendor Risk Management 8-13

8.6.5 Considerations for Small and Mid-Sized Businesses (SMBs) 8-14

8.6.6 Consideration of Distributed Ledger Technology (DLT) in Supply Chain: 8-15

8.7 Supply Chain Industry Standards: 8-16

8.7.1 NIST Cyber Security Framework 1.1 Guidance on Supply Chain Risk Management 8-16

8.8 NIST SP 800-161; Supply Chain Risk Management Practices for Federal Information Systems and Organizations: 8-17

8.8.1 International Organization for Standardization Supply Chain Standards: 8-18

8.8.2 ISO 27000 series; Information Security Standards 8-18

8.8.3 ISO 28000 Series; Supply Chain Security Management: 8-18

9 Conclusions 9-1

9.1 Topic of Report 9-1

9.2 Method of Data Gathering 9-1

9.3 Approach to Analysis 9-2

9.4 Basis of Recommendations 9-2

10 Final CSRIC Recommendations for the FCC 1

10.1 General Recommendations 1

10.1.1 Extension of CSRIC 5G Study 1

10.1.2 Threat Taxonomy 1

10.1.3 Participate with DHS in Public/Private Partnership for Threat Assessment 1

10.2 Recommendations for IoT 1

10.2.1 Regulatory actions 1

10.2.2 Public/Private Partnership 2

10.3 Recommendations for NFV/SDN 2

10.3.1 Regulatory Action 2

10.4 Recommendations for Open Source 2

10.4.1 Regulatory Action 2

11 Final CSRIC Recommendations for Industry 3

11.1 General Recommendations 3

11.1.1 Protecting the 4G Core 3

11.1.2 REST API Authorization 3

11.1.3 REST API Authentication 3

11.1.4 Protection of the User Data 3

11.1.5 Protecting RRC and NAS Signaling 3

11.2 Recommendations for IoT 3

11.2.1 Industry Collaboration 3

11.2.2 NB-IoT Management 4

11.2.3 IoT Best Practices 4

11.2.4 CSRIC IV, WG-5 Remediation of Server-Based DDoS Attacks 4

11.2.5 CSRIC III, WG-7 Botnet Remediation and the Anti-Botnet Code of Conduct 4

11.2.6 CSRIC II, WG-8 ISP Network Protection Practices 4

11.2.7 Comm Sector Coordinating Council White Paper 4

11.3 Recommendations for NFV/SDN 4

11.3.1 Monitor, Monitor and Then Analyze 4

11.3.2 Adopt Industry Driven Standards 4

11.3.3 Future CSRIC Study on NFV/SDN 4

11.4 Recommendations for Open Source 5

11.4.1 Inventory Management of Open Source Components 5

11.4.2 Code Management Systems 5

11.4.3 Testing of Open Source Code 5

11.4.4 Security Frameworks 5

Appendix A – Acronyms 6

# Executive Summary

The telecommunications industry is preparing for the evolution of wireless networks to the next generation of technology, known as 5G. This 5th generation of wireless networks represents perhaps the largest change we have seen in wireless networks since cellular was introduced.

The migration away from traditional, engineered systems designed to support specific network functions in a point-to-point network architecture is moving to adopt an IT architecture. As telecom networks are move into the data center, the future architecture uses IT technologies that have supported the Internet for many years.

This brings about some challenges. This new architecture introduces many new attack vectors that the industry has not experienced before at the same time there is a sense of urgency among operators to complete the standards and begin implementation as quickly as possible. Operators around the world are gearing up for cloud deployments of their critical infrastructure.

The 3rd Generation Partnership Project (3GPP), as well as several other standards organizations, continues work on 5G standards. While many have been focusing on the implementation of 5G networks in 2019, the reality is 5G radio will be deployed on an 4G Long Term Evolution (LTE) network core for these early implementations. A full 5G network core will not be likely until after 2020, and even then, on a limited basis. Wide scale deployment of 5G network core will take time to implement, as we have seen with other technologies (including 4G).

The 5G standards address a number of vulnerabilities that have impacted previous network evolutions. 5G will be more resilient, and introduces new security procedures, however 5G does introduce new attack vectors, and requires a solid security framework to be implemented throughout the network. Fortunately, security considerations are an integral design consideration in the development of the 5G architecture specifications and standards that will result in a more resilient and secure framework for 5G network implementation.

The Communications Security, Reliability and Interoperability Council (CSRIC) VI, Working Group 3 (WG3) has focused on four main areas of 5G technology; Internet of Things (IoT), Network Function Virtualization (NFV)/Software Defined Network (SDN), and open source software,. Maintaining the focus on 5G was critical to meet the stringent deadline for this report. While there are vulnerabilities in these areas in 3G and 4G (studied in previous CSRIC reports), the charter of this working group was to study the risks and vulnerabilities in 5G networks only.

This report focuses on the core network technology of 5G: network security specifications from 3GPP, IoT connectivity, and NFV/SDN. In addition, the report examines the threats and benefits introduced using open source software, which is found even today in networks.

# Introduction

As 5G standards continue to mature, and operators begin planning their deployments of 5G networks, security is once again top-of-mind. The threats identified in prior technologies are still applicable to 5G. 5G will also introduce new attack vectors and a new network architecture that is cloud-based. IoT, SDNs, and open source software will also be a critical ingredient in the 5G ecosystem. This expanded threat environment is being addressed by industry up front through the incorporation of security improvements early in the development of the architecture and specifications for 5G.

The CSRIC WG3 has been tasked to evaluate mechanisms to best design and deploy 5G networks to mitigate risks to network reliability and security posed by the proliferation of IoT devices and open source platforms used in 5G networks. Specifically, the Working Group was tasked to evaluate security risks within:

* IoT
* Open-Source 5G software (e.g. OpenStack)
* NFV and SDN

As part of this assessment, WG3 developed risk impacts for each of the key areas and provided recommendations to mitigate the identified risks and best practices within design, deployment and operation of risk-tolerant 5G.

## CSRIC Structure

|  |  |  |
| --- | --- | --- |
| **Communications Security, Reliability and Interoperability Council VI** | | |
| **Working Group 1: Transition Path to NG911** | **Working Group 2: Comprehensive Re-imagining of Emergency Alerting** | **Working Group 3: Network Reliability and Security Risk Reduction** |
| ***Chair*:** Mary Boyd, West Safety Services | ***Chair*:** Farrokh Khatibi, Qualcomm | ***Chair*:** Travis Russell, Oracle |
| ***FCC Liaisons*:** Tim May, John Healy | ***FCC Liaisons*:** Steven Carpenter, Austin Randazzo | ***FCC Liaisons*:** Vern Mosley, Suzon Cameron |

Table ‑: CSRIC VI Structure

## Working Group 3 Team Members

Working Group 3 consists of the members listed below.

| Name | Company |
| --- | --- |
| Chair Travis Russell, Director, Cyber Security | Oracle Communications |
| Shirley Bloomfield, CEO | NTCA–The Rural Broadband Association |
| Don Brittingham, VP, Public Safety Policy | Verizon Communications |
| Charlotte Field, SVP, Application Platform Operations | Charter Communications |
| Bob Gessner, Chairman | American Cable Association |
| Michael Iwanoff, SVP and CISO | iconectiv |
| Mohammad Khaled, Senior Security Specialist | Nokia Bell Labs |
| Jason Livingood, VP, Technology Policy & Standards | Comcast Cable |
| Jennifer A. Manner, SVP of Regulatory Affairs | EchoStar/Hughes |
| Robert Mayer, VP – Industry and State Affairs | USTelecom |
| Susan Miller, President & CEO | Alliance for Telecom Industry Solutions (ATIS) |
| Drew Morin, Director, Federal Cyber Security Technology and Engineering Programs | T-Mobile |
| Sara Mosley, Acting CTO, OCC/NPP\* | Department of Homeland Security |
| Greg Schumacher, Technology Development Strategist | Sprint Corporation |
| Lee Thibaudeau, CTO & VP of Engineering | Nsight |
| Tim Walden, SVP of Engineering and Construction | CenturyLink |
| Jeremy Larson, Network Manager | USConnect |
| Martin Dolly, Lead Member of Technical Staff | AT&T Services Inc. |
| John A. Marinho, VP Technology & Cybersecurity | CTIA |

Table 2‑2 : List of Working Group Members

The working group heard from several subject matter experts during their research. These Subject Matter Experts (SME) were selected based on their area of expertise in 5G architecture and technology. They brought a different perspective for the working group members to consider when developing their recommendations.

**SMEs, Acknowledgements:**

* Eric McMurry, Dir., Software Development, Office of the CTO, Oracle Communications
* Roger Piqueras Jover, Wireless Engineer, Senior Security Architect, & Security Research Scientist at Bloomberg L.P.
* Peter Schneider, Senior Security Researcher, Nokia Bell Labs

# Introduction to 5G

5G is the commonly used acronym assigned to the 5th generation of wireless technology that has both a fixed and a mobile component. It is much broader than previous generations of wireless networks. While previous generations of wireless services were focused on a single type of network (i.e., mobile wireless), 5G has been defined as a “network of networks.” This is because of the recognition that a single technology (e.g., mobile wireless) will not be able to provide all the characteristics of 5G such as ubiquity and resiliency.[[1]](#footnote-1)

The CSRIC VI WG3 was asked to examine the security risks associated with this new technology that will require both infrastructure and device changes that may introduce incremental security risk.

## Who Is Defining 5G?

The 3rd Generation Partnership Project (3GPP) unites seven telecommunications standard development organizations (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC), known as “Organizational Partners”, and provides their members with a stable environment to produce the reports and specifications that define 3GPP technologies[[2]](#footnote-2). 3GPP standards work comes in releases. Currently, release 15 defines the first set of 5G requirements, including continued work on the definition of LTE-Advanced Pro. LTE-Advanced is the current 4G network core that many network operators will be launching 5G radio with, until the full work of the 5G core is complete. There are a number of standards bodies that will work to enable 5G.

No single standards body can be said to control the definition of 5G. Of course, 3GPP, which has been instrumental in the development of mobile wireless standards to date, will continue to play a leading role, especially as it evolves to include non-terrestrial technologies such as satellite and high-altitude platforms. In addition, other standards bodies, such as ATIS and GSMA will play an important role.

The ITU 5G Infrastructure Public Private Partnership (5GPPP) has been working with 3GPP and others to define spectrum requirements, as well as other areas not addressed by 3GPP. Likewise, the Internet Engineering Task Force (IETF), the Next Generation Mobile Networks (NGMN) Alliance, IEEE, the GSM Association (GSMA), and Small Cell Forum have been contributing requirements.[[3]](#footnote-3)

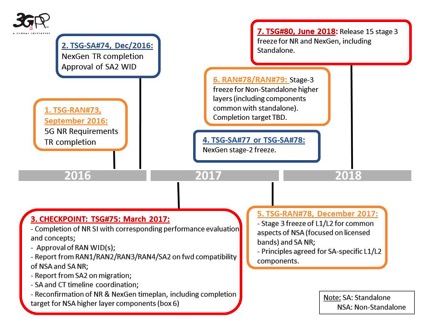


Figure ‑: The above figure shows the timeline for 3GPP release 15. The start of release 16 is also depicted in this chart. SOURCE: 3GPP

### Status of 5G Standards

The 3GPP standards work comes in releases. Currently, release 15 defines the initial set of 5G specifications and requirements, including continued work on the definition of LTE-Advanced Pro. There are many gaps in R15 that need to be filled in later releases. For example, R15 does not address congestion management, traffic management, or routing management in the service based architecture (SBA). This was managed by a Diameter Routing Agent (DRA) in 4G, and the Signal Transfer Part (STP) in SS7. There are proposals being submitted to 3GPP now to define another function that would provide this capability, but this will not be defined until after R15. This means that 5G standards are not currently at a stage where a 5G core network can be deployed with any guarantee of resilience and availability.

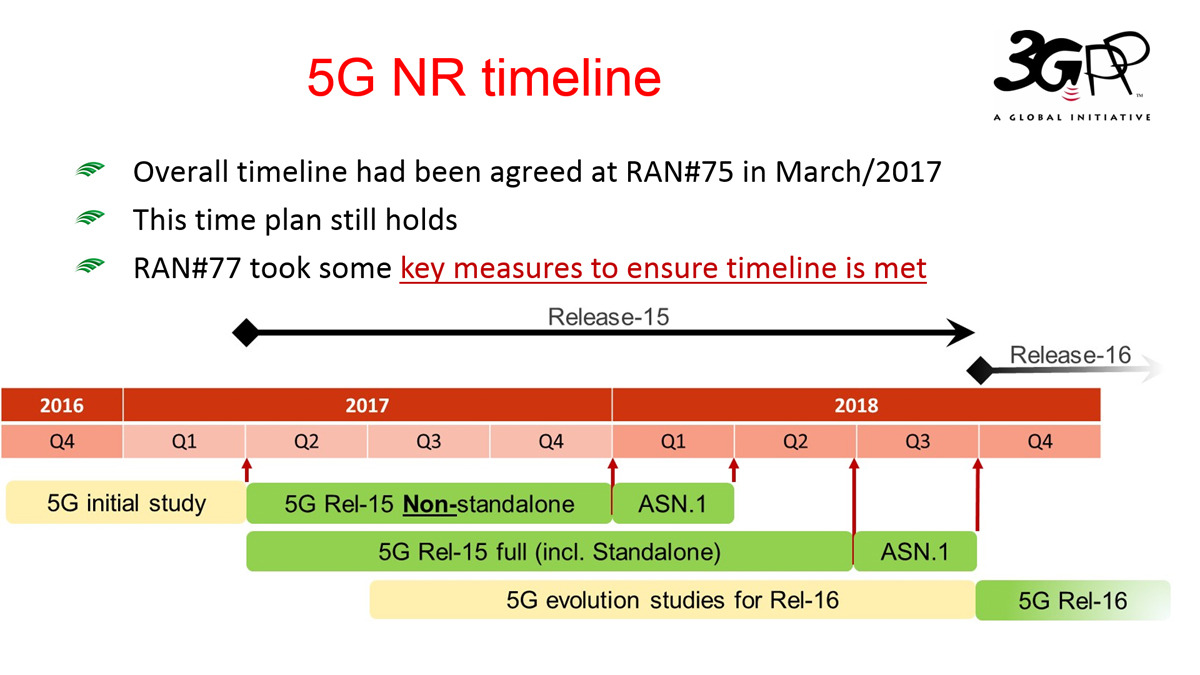


Figure ‑: The 3GPP release timeline for R15

Release 16, considered “5G phase 2,” is scheduled to be completed in December 2019. This release includes finalization of the 5G core architecture and interfaces, as well as defining non-terrestrial solutions for access. Standards development is well under way for both terrestrial and non-terrestrial technologies. The figure above shows the work from the Standardization Special Interest Group (SSIG) focusing on non-terrestrial networks in 5G.

## The 5G Mission

Key performance objectives of 5G are defined in terms of Enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low-Latency Communications (URLLC), and massive Machine Type Communications (mMTC). While much focus has been on the improvements in data rates over what exists today, the reductions in response times/latency and corresponding connection density supported by 5G enables the transformation of a number of industries. For example, this improved latency enables real time capabilities to allow autonomous cars and the ability to have traffic devices control the flow of traffic. In addition, 5G supports denser connectivity for IoT devices supporting both low power and high-power devices. This is enabled by improvements in spectral efficiency that supports the massive capacity requirements for this transformative technology.

| 5G Performance Requirement Type | Minimum KPI requirement | Category |
| --- | --- | --- |
| Peak Data Rate | Downlink: 20 Gbps  Uplink: 10 Gbps | eMBB |
| Peak Spectral Efficiency | Downlink: 30 bits/sec/Hz  Uplink: 15 bits/sec/Hz | eMBB |
| Data rate experienced by User | Downlink: 100 Mbps  Uplink: 50 Mbps | eMBB |
| Area Traffic Capacity | Downlink: 10 Mbits/sec/m2 in indoor hotspot (eMBB test environment) | eMBB |
| Latency (User Plane) | * 4 ms for eMBB * 1 ms for URLLC | eMBB, URLLC |
| Latency (User Plane) | * 20 ms  (10 ms encouraged) | eMBB, URLLC |
| Connection Density | 1 x 106 devices/Km2 | mMTC |
| Average Spectral Efficiency | (All the below figures are in units of bits/sec/Hz/TRxP)  Indoor hotspot: DL:9/ UL:6.75  Dense Urban: DL: 7.8/ UL: 5.4  Rural: DL: 3.3/UL: 1.6 | eMBB |
| Energy Efficiency | * Efficient data transmission (Loaded case): To be demonstrated by "average spectral efficiency". * Low energy consumption (no data case): This test case should support high sleep ratio/long sleep duration. | eMBB |
| Reliability | 1 x 10-5 probability of transmitting layer-2 PDU of 32 bytes in size within 1 ms (in channel quality of coverage edge for Urban Macro-URLLC test environment.) | URLLC |
| Mobility | * Dense Urban: up to 30 Km/h * Rural: up to 500 Km/h | eMBB |
| Mobility Interruption Time | 0 ms | eMBB, URLLC |
| Bandwidth (Maximum Aggregated System) | * At least 100 MHz * Up to 1 GHz for operation in high frequency bands i.e. above 6 GHz | IMT-2020 |

Table ‑: 5G Key Performance Indicators. SOURCE: RF Wireless World[[4]](#footnote-4)

Over the years, wireless networks have continued to rely on the same basic architecture. When data was added to mobile networks, an overlay was used for the routing of data packets. Even with the 4G network architecture, data continues to be forced through the same architecture used since 3G.

This traditional 3G architecture does not serve the connection of billions of devices as forecasted for IoT. Nor does this architecture work well for supporting the video sessions forecast to consume our mobile networks. Indeed, Cisco estimates that by 2020, 75% of all mobile data will be video.[[5]](#footnote-5)

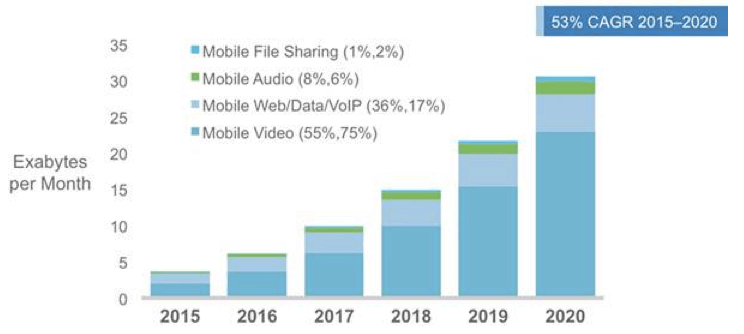


Figure ‑: Cisco’s VNI report shows video representing the bulk of mobile data by 2020. SOURCE: Cisco

To better support video and other mobile data sessions, and to eliminate the need to build huge pipes to transit all this data, 3GPP set out to remove what is not needed for every single data session. In an effort to reduce latency in the network to less than 5 milliseconds, the network architecture has to change.

In addition to latency, the radio infrastructure has to be redesigned as well. Connecting billions of devices is not possible with the current radio access network. Antennas and radios need to be brought in much closer to the device, and spectrum management must be more efficient to support all this connectivity. These were the two priorities for developing 5G radio requirements.

### Speeds and Connectivity in 5G

Each generation of wireless technologies builds on the capabilities of the prior generation, and the most tangible change is the significant increase in network speed. 5G will be no exception. 4G LTE networks are fast, with speeds that can approach ~100 Mb/s peak rates. But thanks to innovations like millimeter (mm) wave technology, 5G networks are expected to generate a throughput more than 10 times faster, particularly in dense urban locations and initially likely in localized ‘hot spots.’ 5G speeds more than 1 Gb/s have been demonstrated in initial testing.

4G or the 4th Generation Mobile offered speeds of up to 100Mb/s for users on the move. 4G devices allow users to take advantage of the high-speed networks that have been built, however 4G radio access is not ubiquitous throughout the US due to the high capital cost and the distribution of the population.

For 5G to be truly successful, it will have to be ubiquitous and reliable. With regard to ubiquity, there are always going to be portions of the globe where it is just too expensive or impractical for terrestrial wireless networks to reach. For these areas, satellite has been and will continue to be the answer, as might other non-terrestrial technologies, such as high-altitude platforms (HAPS). For example, when wide area networks are required, satellite may prove to be a better option.

5G is expected to enable entirely new services and applications with latency values of 6 to 9 msec – far less than current 4G networks. However, some believe that we will require latency of 1 msec to support capabilities such as an autonomous car to be able to deal with the quick response times needed. 4G LTE today has latency in a range of 30 to 45 msec, so even at the demonstration level of 1 to 10 msec 5G is showing today, this is a significant improvement. Clearly, the current latency associated with 4G is more than sufficient to support Voice over IP (VoIP) or video calls with minimal delay or jitter. However, the improvement with 5G will enable more real time applications that customers will find of significant benefit.

5G low latency minimizes delays introduced through network functions enabling new services and applications. Edge computing pushes the processing of data to the edge of the network where it can be processed much quicker.

To put latency into context, 4G latency rates – the technical term for the delay between your request for data and when your mobile device receives it – are low, roughly 10 milliseconds over-the-air; 50 milliseconds end-to-end. This is what enables good quality VoIP calls and video calls with little delay or jitter. 5G latency is expected to be 5 to 10 times lower than 4G.

Multiple technologies, including HAPS and non-geostationary orbit satellite systems will be able to meet these latency requirements. When high capacity is needed, and latency is not an issue, high-capacity geostationary satellite systems will play an important role.

But 5G is much more than just faster speeds and lower latency. With this new technology, 5G provides the scale for wireless networks to support far more devices and applications than is possible in 3G or 4G networks. Specifically, the massive number of new wireless IoT devices – nearly 28 billion in total and an estimated 1.5B M2M and consumer connected wireless (cellular) devices by 2021- will rely on 5G.

### Applications Supported by 5G

There will be many applications and services that are enabled by 5G, many of which will require a more secure network connection than what is available today. This is especially true in industry sectors such as healthcare and smart cities. However, this flexibility introduces complexity as well.

To address this complexity, 5G system architectures will leverage a SOA (Service-Oriented Architecture)[[6]](#footnote-6) approach to decompose complex systems into applications consisting of a system of simple well-deﬁned components. In SOA architecture, each system oﬀers its functionality as standard services typically through a web services application-programming interface (API). The figure below depicts a general-purpose SOA middleware architecture composed of the following component layers:

* **Applications:** This layer enables end users to request information services and inter-act with the middleware using an API.
* **API:** Applications implement the methods provided by this layer to access services.
* **Service Provision:** Each available service has a respective infrastructure of devices connected to the middleware, so the devices function is abstracted into a service and provided in this layer.
* **Management Core:** This is the main layer of the middleware. It is composed of functions that allow the management of each device in the environment.
* **Devices Abstraction:** This is the lowest layer of the middleware and facilitates the management of the devices through a common language and procedures.
* **Devices:** This layer is composed of any device. These devices can connect to the middleware and provide data to the upper layers.
* **Security:** The support of security is an important feature of any middleware system. The middleware must be able to provide functions related to security for all exchanged and stored data. These functions can be either built on a single layer or distributed among all layers.

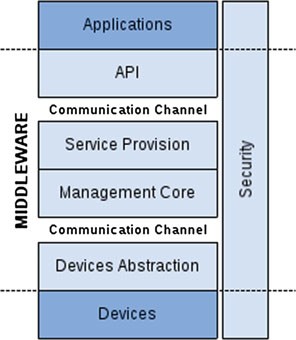


Figure ‑: The service oriented architecture introduced in 5G eliminates the need for massive data pipes. SOURCE: ResearchGate White Paper, Figure 1, page 405

The 5G network promises to be much more complex than its predecessors, likely using a heterogeneous architecture comprised of multiple access technologies. There are multiple stakeholders with different views and requirements of secure networks:

* Network Operators, who need to protect revenues, their brand, SLAs with their customers, and providing data protection
* National Motivations, including national security, protection of infrastructure and development of national/global economies.
* Enterprise and consumers, who are expecting high levels of privacy for their personal data
* Service providers, who require highly reliable and resilient networks, strong end user authentication, and access controls to ensure the availability of products and services for end consumers

Service providers[[7]](#footnote-7) will be expected to meet the needs of both consumers and national motivations, while supporting new content delivery through a carrier-grade network with high availability. The principal motivations of the network operator include revenue protection, brand protection (which includes protecting consumer data to prevent damage to their own brand), meeting license conditions and compliance, and offering new services to their customers (such as adult content filtering and digital rights management). They need to maintain their “carrier-grade” differentiation for commercial reasons. Note that in many cases, network operators will also be service operators providing not just the connectivity, but also end user products and services.

Nations, and their various government agencies, have a different set of motivations. This includes the protection of the national infrastructure and the development of national and global economies. This includes support for:

* Emergency Services Networks with particular requirements in disaster situations, including Citizen to Authority (911 services and FirstNet)
* Authority to Citizen (Public Warning Systems) and Authority to Authority (GETS, WPS)
* Support to law enforcement such as Lawful Interception and Retained Data
* Protection of Critical Infrastructure (CI). Communications are increasingly integral to most other aspects of CI such as water and power. 5G networks will also create and underpin entirely new components of critical infrastructure, such as tactile internet or remote monitors in health care and vehicle communications in transport
* Defense of national/global economies to provide confidence to do business securely, to maintain an on-line global economy, and to maintain the availability of communications as a core foundation of global trade.

### The Internet of Things

The IoT will enable connectivity between and among networks. It introduces an opportunity where everything we interact with is interconnected and communicating with many other devices. Our lives may be made more convenient not just through appliances and gadgets, but also through smart meters, medical devices, industrial controls, and even self-driving cars. But aside from convenience, IoT can also provide lifesaving services through connected sensors and devices.

It is estimated that billions of devices will be connected to the network by 2020. Many of these devices will be low power sensors, wearables, and small devices used in industry. With a greater density in radio access, and with the radio moving closer to the device, low power devices are now possible.

Other IoT uses require only low data rates and not time sensitive communications. There are a number of different IoT uses foreseen, each with a unique set of network requirements:

* Wearables
* Industrial IoT
* Smart Homes
* Connected vehicles
* Smart Cities
* M-Health and Telemedicine

This means that these networks will need to be interoperable in many different ways.

Another critical role for 5G networks is providing connectivity to ensure that patients and other vulnerable populations, such as the elderly, including those who live in rural and remote locations, are connected. 5G will be able to provide critical connections to these populations.

## 5G Architecture

5G represents a major change in network architecture. Previous generations of networks used a hub and spoke model. In 5G, the control plane is a bus architecture. The network functions on the bus are only used by services that need the functionality, making this a SOA (3GPP has declared this an SBA).



Figure ‑: The 5G architecture when compared to 4G EPC. SOURCE: Oracle Communications.

In 4G, there were only about 6 or 7 network functions that were defined. In 5G, many of these network functions are broken apart into multiple functions. There are now more than 27 network functions defined, and more are being defined in subsequent releases. Following are some of the network functions defined for 5G:

* **AMF:** This is the Access and Mobility Management Function. This replaces the Mobility Management Entity (MME), which was split into two different functions.
* **AUSF:** The Authentication Server Function. This provides the authentication services found in the 4G MME.
* **UPF:** The User Plane Function. This is part of the S-Gwy/P-Gwy function found in 4G and is used for managing the user plane data.
* **SMF:** The Session Management Function. This is the other half of the S-Gwy/P-Gwy functionality in 4G. The SMF manages the sessions supported by the UPF.
* **UDM:** The Unified Data Management. This is what was the Home Subscriber Server (HSS) in 4G networks.
* **PCF:** The Policy Control Function. This is the same as the Policy and Charging Rules Function (PCRF) in 4G.
* **AF:** The Application Function. This is similar to the Proxy-Call Session Control Function (P-CSCF) used in IP Multimedia Subsystem (IMS) networks.
* **NEF:** The Network Exposure Function. This provides some of the capabilities of the Service Capability Exposure Function (SCEF) defined in 4G.
* **NSSF:** The Network Slice Selection Function. This is a new function defined for managing selection of network slices.
* **NRF:** The Network Repository Function. This is also a new function defined for 5G. The NRF is used for discovering and registering networks and the services they provide.
* **NWDAF:** The Network Data Analytics Function. This is a new function defined for 5G for collecting KPIs from the various network functions for analytics.
* **SEPP:** The Security Edge Protection Proxy. This is a new function defined for 5G to protect the network edge, and act as the gateway between the home network and visited networks.

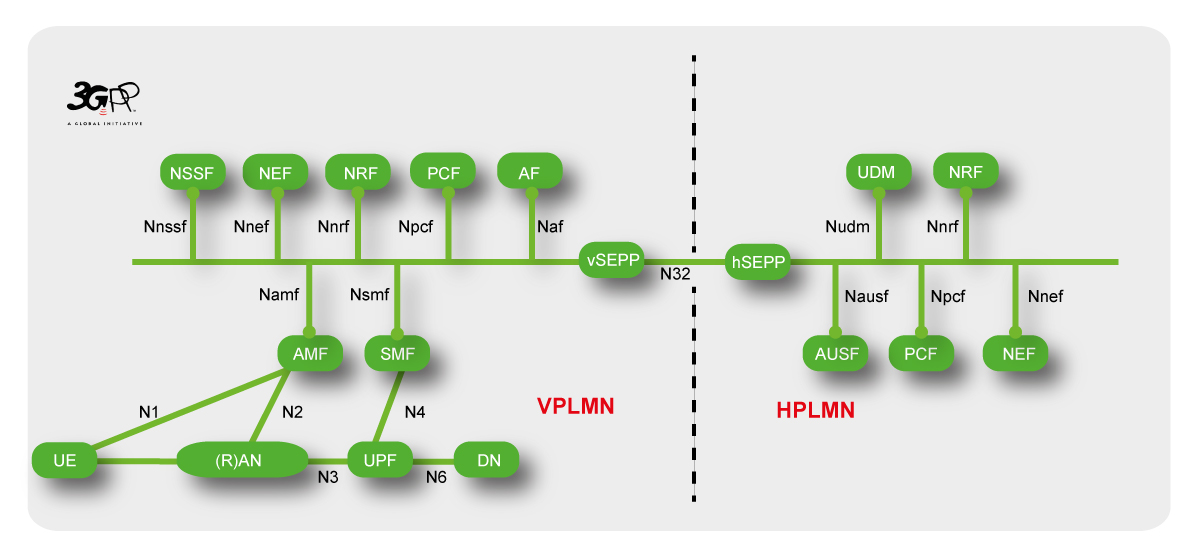


Figure ‑: Interconnecting multiple networks via 5G uses the N32 interface, depicted here, with a new security function, the Security Edge Protection Proxy (SEPP). SOURCE: The 3GPP

5G is being deployed in two different phases. The 3GPP has defined 11 different options for implementing 5G in its early stages, allowing operators to use their existing investment in 4G Evolved Packet Core (EPC) with 5G radio. Out of these 11 different options, operators are seriously considering only a few.

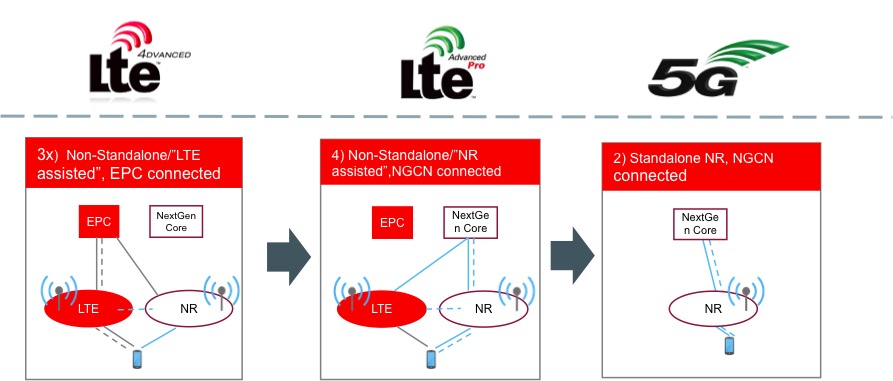


Figure ‑: The three phases of 5G implementation. SOURCE: Oracle

The first option is known as 3x, and is the most widely selected option. The operators that have announced 5G services in 2018 are using this option to deploy 5G radio using their 4G EPC. The signaling from the device to the network core uses traditional 4G signaling but introduces the ability to add 5G radio connectivity. Operators will then migrate from 3x to 7x, where they are introducing the 5G network core, but still supporting the 4G radio. The 4G EPC is no longer used at this point. This represents the largest migration risk, because the new network core is being implemented for the first time. Some point later in time, the 5G core and 5G new radio will be deployed as a standalone network without any 4G components.

The protocols to be used in 5G are evolving as well. Initially, Hypertext Transfer Protocol (HTTP)/2 will be used on Transmission Control Protocol (TCP) transport in the control plane. The signaling will be JSON, using REST APIs. However, there is already discussion of migrating from TCP to Quick UDP Internet Connections (QUIC). QUIC runs on UDP, which provides faster connections and less latency, while still supporting the connection-oriented features of TCP. HTTP/2 will ride on QUIC (see table below).

|  |
| --- |
| REST |
| JSON |
| HTTP/2 |
| QUIC |
| UDP |
| IP |

Table 3‑2: 5G protocol stack. SOURCE: Oracle Communications

There are other protocols used in 5G. The GPRS Tunneling Protocol – User Plane (GTP-U) is used to manage the user plane data sessions from the Radio Access Network (RAN) to the UPF. The Non-Access Stratum (NAS) protocol is used between the RAN and AMF. Both of these are existing protocols found in 4G networks.

### 5G New Radio (NR)

5G includes a new flexible radio interface (NR=New Radio) along with a new scalable and flexible 5G Core network, designed to allow new features and capabilities. Both the Radio Access and Core Networks have significant changes from 4G networks and were designed to support a broad range of use cases including: enhanced mobile broadband (higher data rates and lower latency), massive IoT (ultra-high density of devices, ultra-low energy) and mission critical services (ultra-high reliability, low latency and strong security). To accomplish this, NR was designed from the ground up to be flexible to meet these needs. A few of the specific enhancements supported by NR include:

* Operation in the mmWave bands. Today’s LTE networks use sub 6 GHz spectrum, but NR is designed to operate in both sub 6 GHz and mmWave spectrum. To efficiently operate across such a broad range of spectrum, NR requires a flexible frame structure using OFDM with different sub-carrier spacing and operational with multiple bandwidths (e.g. 5 – 100 MHz bandwidths for sub 6 GHz bands and up to 400 MHz bandwidths for the mmWave bands)
* To reduce the latency, NR allows for scalable slot durations, mini-slots, a self-contained slot structure and traffic pre-emption for Ultra-Reliable Low Latency Communication, and other new capabilities
* To support the higher data rates, NR will leverage the higher bandwidths often found in mmWave bands.
* NR also utilizes new forward error correction schemes such as LDPC (Low Density Parity Check) and Polar Coding. These techniques improve throughput, reliability and also reduce complexity of the radio.
* NR will be a “beam based” air interface that supports advanced wireless technologies such as Massive Multiple Input Multiple Output (MIMO). Massive MIMO is the extension of the traditional MIMO technology to antenna arrays having a large number (>>8) of controllable antennas. In Massive MIMO, transmission signals from the antennas are adaptable by the physical layer via gain or phase control. Massive MIMO will enhance coverage in 5G NR by using high gain adaptive beam forming and will enhance capacity using high order spatial multiplexing.

5G NR will utilize multiple spectrum types (licensed, unlicensed, shared) and flexible spectrum coverage footprint with a combination of low, mid and high bands. High bands will mainly be used for higher capacity while low bands will be used for IoT applications and low latency critical communications.

Many early deployments will leverage the existing radio equipment, including the legacy 4G LTE core network and the legacy LTE radio network. 3GPP defined several deployment options to allow operators to smoothly migrate from their 4G network to a 5G network.

5G implementations are most likely to follow three phases; phase 1 relies on 4G LTE EPC, with 4G and 5G radio. Phase 2 includes both the 4G and 5G RAN network on a 5G core. The final phase will consist of 5G RAN and core.

The 5G base station (gNB) has been split into two functions: the central unit and the distributed unit. This allows for much higher data rates with lower latency. Currently, 3GPP standards have defined 5G NR as supporting both high frequency (>6 GHz) and low frequency (<6 GHz) bands. This ensures that 5G NR will support a wide variety of services for a wide range of different devices.

Satellite is also being considered as part of the NR. The objective of TSG RAN TR 38.811 “Study on NR to support non-terrestrial networks” is to study channel models, define the deployment scenarios as well as the related system parameters, and to assess potential key impact areas on the NR. Specifically, 3GPP is ensuring that satellite network elements are fully compatible with 5G protocols, so that use of non-terrestrial networks is transparent to 5G.

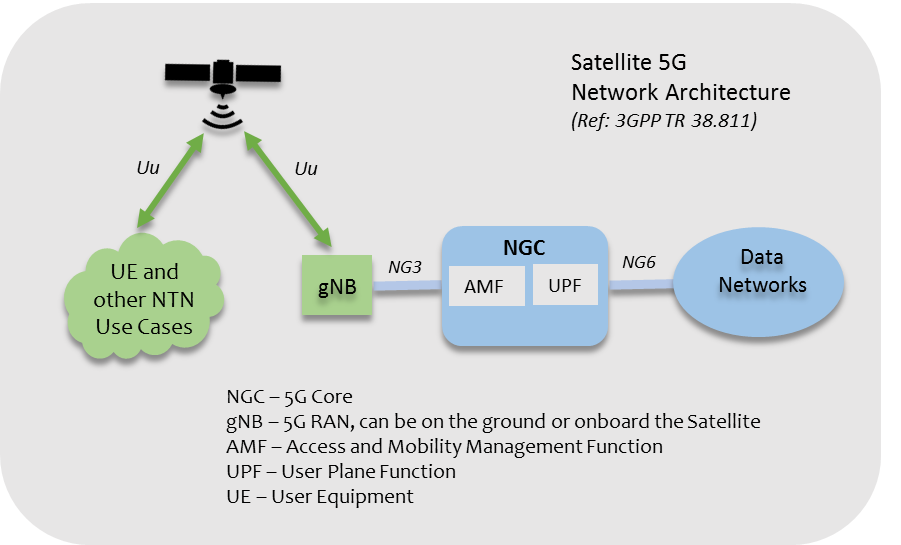


Figure ‑: Satellite radio considerations for 5G. Source: 3GPP TR 38.811

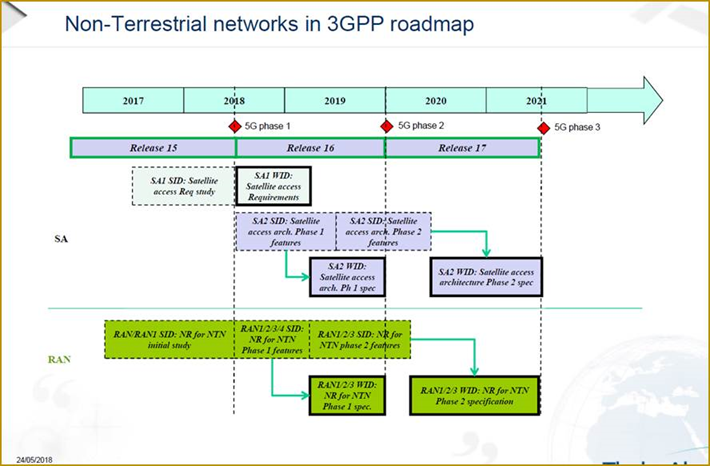


Figure ‑: Timeline for requirements for non-terrestrial networks to be used in 5G implementations. SOURCE: 4th SSIG (Standardization Special Interest Group) meeting May 24, 2018

#### Threats introduced through the 5G NR Threat Surface:

It is expected that hackers will continue to target 5G networks as they did 3G and 4G. Location tracking, traffic interception, network spoofing, impersonation of mobile devices, malicious use of fake base stations, Denial of Service (DoS) attacks using various forms of jamming, including smart jamming where DoS in a cell is achieved with very limited transmission power by targeted jamming of narrow but essential control channels; these are all going to be attempted in 5G networks. However, there have been a number of improvements in 5G that will mitigate the effects of these attacks, if not prevent them altogether.

In general, NR doesn’t raise the need for higher radio security or higher jamming protection than it is in LTE. Rather, the new wide variety of services and applications for wide range of different devices may require this.

For example, IoT devices for massive deployments may require security algorithms and procedures that provide high security despite a low overall energy budget at the device. In this context, physical layer security may also need to be considered, where the network may authenticate the origin of a message sent by a device by verifying the known characteristics of the specific transmission channel from the device, or other characteristics of the device that cannot be imitated by an attacker. Use of such “physical unclonable functions” may relieve the device from the need to compute a cryptographic message authentication code. On the other hand, mission critical services including military and public safety communications may require higher resistance against DoS attacks against the network, including better jamming protection.

While not specifically related to the radio technology, new deployment options for radio equipment may raise new threats to 5G radio networks. When a 5G base station (gNB) is deployed similar to a typical LTE base station (eNB), i.e. as a stand-alone box comprising all RAN functions, threats involving physical tampering are applicable in a very similar way. However, gNBs may also be deployed based on edge clouds, with virtualized network functions running on infrastructure shared with other software applications. This raises new threats, such as side channel attacks via shared infrastructure. Although these are not NR specific, these attacks apply in general to cloud based networks, including 5G core networks. Also, the physical security of edge cloud deployments may be a concern. Integrity protection mechanisms for the cloud software platform and the virtual network functions need to be considered to detect possible compromise by attackers with physical or remote access.

According to current 3GPP specifications, 5G will introduce some security enhancements, e.g. better protection of the subscriber privacy by preventing “International Mobile Subscriber Identity (IMSI) catching.” It is therefore reasonable to look for ways to prevent “bidding down” attacks that force a device attachment to earlier mobile network generations and ensure that the most secure technology is used wherever it is available.

#### Spectrum Sharing Threats surface

5G NR will operate in a flexible frequency range. In addition to operating in licensed and unlicensed spectrum environments, 5G will also operate in a shared spectrum environment. In the United States, several lower, mid and high bands are expected to be available for spectrum sharing in 5G. 3GPP will likely have a new 5G band dedicated for the United States Citizen Band Radio Service (CBRS) that will use a shared wireless broadband. The erroneous management of the spectrum in a shared environment increases the risk for interferences and is governed by a three-tiered spectrum authorization framework. New spectrum sharing access and operations will be managed by Spectrum Access Systems (SASes). SASes are required to employ protocols and procedures to ensure that all communications and interactions between the SAS and Citizens Broadband Radio Service Devices (CBSDs) (5G gNBs in the 5G context) are accurate and secure and that unauthorized parties cannot access or alter the SAS or the list of frequencies sent to a CBSD. A compromised SAS under the control of an adversary might drastically alter the communications between SASs and CBSDs and potentially disrupt operation of all CBSDs (5G gNBs) under that SAS.

### Multi-access Edge Computing (MEC)[[8]](#footnote-8)

Multi-access Edge Computing (MEC), formerly Mobile Edge Computing, is a network architecture that offers application developers and content providers cloud-computing capabilities and an IT service environment at the edge of the network. The basic idea behind MEC is that by running applications and performing related processing tasks closer to the end device, network congestion is reduced, and applications perform better.

Operators can open their RAN edge to authorized third parties, allowing them to flexibly and rapidly deploy innovative applications and services towards mobile subscribers, enterprises and vertical segments. This environment is characterized by ultra-low latency and high bandwidth as well as real-time access to radio network information that can be leveraged by applications.

MEC will enable applications and services to be hosted ‘on top’ of the mobile network elements, above the network layer. By moving compute, storage and processing power closer to the network edge, MEC will be key to delivering on 5G use cases for new vertical business segments and services for consumers and enterprise customers.

By deploying various services and caching content at the network edge, mobile core networks are alleviated of further congestion and can efficiently serve local purposes. These applications and services also benefit from being in close proximity to the customer and from receiving local radio-network contextual information. MEC uniquely enables software applications to tap into local content and real-time information about local-access network conditions. Other benefits of MEC include:

* A reduction of backhaul traffic
* Cost reduction by decomposing and disaggregating access function
* Optimization of central office infrastructure
* Improve network reliability by distributing content between the edge and centralized data centers
* Deliver innovative services not possible without edge computing

MEC industry standards and deployment of MEC platforms will act as enablers for new revenue streams to operators, vendors and third parties. Differentiation will be enabled through the unique applications deployed in the Edge Cloud that leverage the MEC APIs, management interfaces and essential platform functionality.

#### MEC Initiative

The MEC initiative is an Industry Specification Group (ISG) within ETSI. The purpose of the ISG is to create a standardized, open environment that will allow the efficient and seamless integration of applications from vendors, service providers, and third parties across multi-vendor Mobile-edge Computing platforms deployed by network operators.

The work of the MEC initiative aims to unite the telco and IT-cloud worlds, providing IT and cloud-computing capabilities within the RAN. The MEC ISG will specify the elements that are required to enable applications to be hosted in a multi-vendor mobile-edge computing environment.

#### MEC Security requirements

The mobile edge system needs to provide a secure environment for running services for the following actors: the user, the network operator, the third-party application provider, the application developer, the content provider, and the platform vendor. The mobile edge platform must also restrict the mobile edge application to access only the information for which the application is authorized. In practice, these security requirements are addressed through proper application of best practices for a SOA and API security as described below in the sections on securing IoT in 5G.

One other consideration for MEC is physical security. Since MEC relies on locating part of the core of the network into remote sites for latency-centric applications essential to certain 5G use cases, it is necessary that the physical security of these assets should be prioritized. In addition to the physical controls and barriers used in current deployments to protect RAN assets, remote monitoring of the MEC will be necessary to improve reliability and resilience from threats.

### Network slicing

Network slicing is a form of virtual network architecture utilizing the principles behind SDN and NFV, based on the separation of the control and user planes. When implemented, it enables the operator to create multiple virtual networks over a common physical network infrastructure. This allows the operator to tailor services to the unique requirements of each individual customer while utilizing network resources most efficiently and preserving the security of services for each customer.

5G mobile networks are expected to support multiple network slices running on a common infrastructure with different services like core network slices, RAN slices, and e2e slices. Slices can use a common infrastructure (NFV infrastructure, SDN-based transport) and may use common functions (e.g. a common radio scheduler may assign radio resources in a cell to different slices).

Each unique virtual network, or network slice, is comprised of an individual set of logical network functions that support the particular use case of the customer. Examples of network slices include private mobile network, enterprise business service, enterprise Wi-Fi, low power / low bandwidth IoT, and low latency AR/VR. Each slice can have its own unique functional requirements as defined by factors such as speed, connectivity, capacity and latency. However, these functional requirements can also be shared across different network slices.

In 3GPP, network slicing is being defined in TS 23.501. A network slice is defined within a Public Land Mobile Network (PLMN) and includes the core network control plane and user plane network functions as well as the 5G access network (AN). The 5G AN may be:

* A next generation RAN described in 3GPP TS 38.300, or
* A non-3GPP access network where the terminal may use any non-3GPP access to reach the 5G core network via a secured IPsec/IKE tunnel terminated on a Non-3GPP Inter-Working Function (N3IWF), as described in 3GPP TS 23.501 and 23.502.

TS 23.501 further defines network function, slice, and slice instance as follows:

* Network Function: A 3GPP adopted or 3GPP defined processing function in a network, which has defined functional behavior and 3GPP defined interfaces. (Note: A network function can be implemented either as a network element on dedicated hardware, as a software instance running on dedicated hardware, or as a virtualized function instantiated on an appropriate platform, e.g. on a cloud infrastructure.)
* Network Slice: A logical network that provides specific network capabilities and network characteristics.
* Network Slice Instance: A set of network function instances and the required resources (e.g. compute, storage and networking resources) that form a deployed network slice.

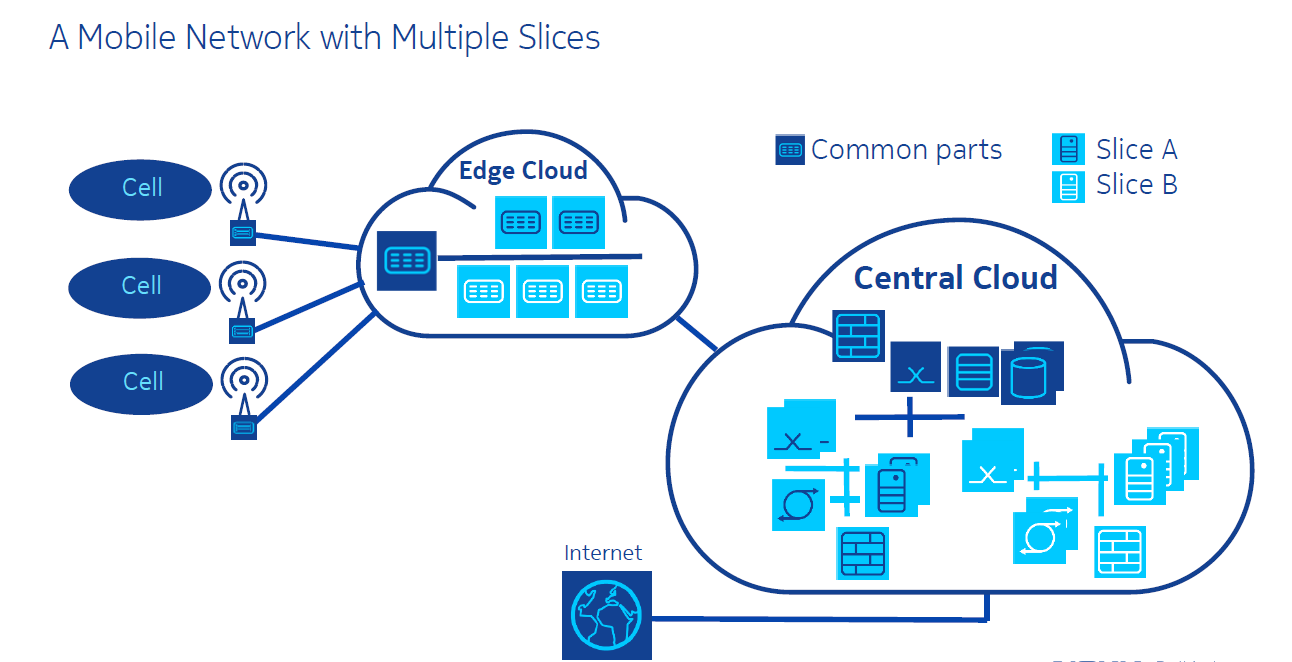


Figure ‑: Network slicing elements. SOURCE: NOKIA

Security and trust aspects are a concern when using shared network infrastructure. Since network slices are a new concept and not yet deployed widely, we have not seen attacks on network slices in the wild. Some potential attacks against network slices may include:

* DoS attacks on “small” slices.
* Attacks on interfaces to common network parts (vertical mobile network operator)
* Attacks on management interfaces provided for verticals to manage their slices
* Attacks via inter-slice interfaces
* Attacks on slicing-specific procedures: slice selection, slicing-specific authentication and authorization, slice management
* Malicious message routing between different slices.

Most of the above attacks can be mitigated by full isolation of the slice against other users of the common network infrastructure.

To achieve full isolation, both resource isolation and security isolation need to be considered. Resource isolation can be achieved by making sure that resources dedicated to one slice cannot be consumed by another slice. Security isolation should make sure that data/traffic cannot be intercepted or faked by entities of another slice.

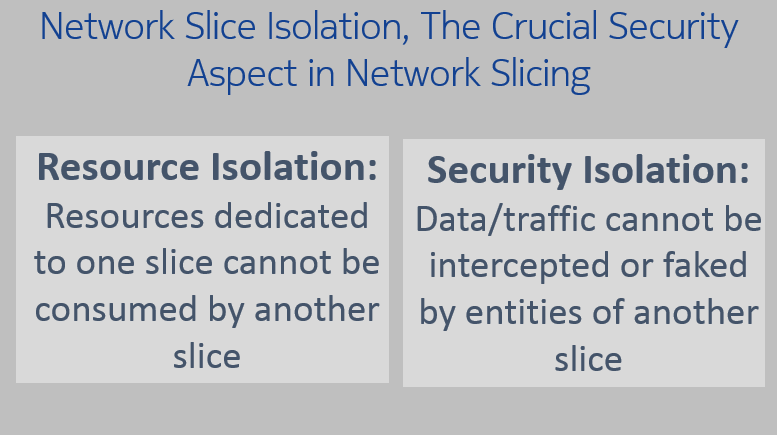


Figure ‑: Security considerations for network slicing. SOURCE: Nokia

In addition, user equipment needs to be authenticated and authorized for accessing the specific slice. Access to the slice is usually indicated using slice specific identifiers by the user equipment to the network signaling. If this signaling is not security protected, access to the slice itself may be denied resulting in a DoS attack. Hence just as any other signaling to the core network, all signaling meant for access to the network slice as well as UE to network slice signaling needs to be protected.

In a 5G telco cloud environment, sharing cloud infrastructure brings the risk of breaches via flaws in the cloud software (e.g. hypervisors). This software must be designed and implemented with highest care to minimize its vulnerability. Isolation between different network slices running on a virtual shared infrastructure requires preventing virtual machines in one slice from impacting those in other slices. It is also required to prevent information from leaking between slices on side channels (e.g., via physical memory sequentially used by different slices). In addition to the advantage of more efficient network resource utilization when compared to traditional network architectures, network slicing, when properly implemented, provides enhanced security by isolating network slices from one another. Unless there is assurance of the isolation of each network slice, the security of the data moving within a slice cannot be trusted.

Network slicing will be especially critical in 5G networks. The immense variety of use cases, and the variability in the performance requirements of each use case, dictates the need for the network efficiency, scalability, and flexibility delivered by network slicing capability. For example, support for the different performance requirements of use cases such as remote surgery, vehicle-to-vehicle communication, and viewing a cat video can all be accommodated on a single physical network infrastructure through the use of network slicing. Additionally, once any of those sessions is complete the network resources can be dynamically recovered and utilized for a new network slice with different performance requirements.

### Cloud and virtualization

NFV/SDN introduces complexity to the network, and many operators may not implement fully orchestrated NFV/SDN. However, they will still want to take advantage of virtualization, and cloud technology. Virtualization does not mean NFV is a requirement; in fact, an operator can virtualize their network without implementing NFV.

The principles behind virtualization are the same as with NFV, without autonomous orchestration. Network functions will have to be instantiated manually. Many operators are using this strategy to cap and grow their current network infrastructure. Rather than purchase more dedicated hardware, they are opting to purchase software-based network functions and deploy them on their own hardware.

The same considerations from a security perspective apply when implementing network virtualization. Isolation of critical network functions is important, and care with the implementation of hypervisors applies here as well.

# 3GPP SA3 (Security) Standards

The 3GPP standards process is an ongoing effort that will continue to release updates for adoption. While not the latest release for comment, the report references 3GPP TS 33.501, Specifications for 5G Security, published by 3GPP SA3. This specification defines the security architecture (i.e., the security features and the security mechanisms for the 5G system and the 5G core) and the security procedures performed within the 5G system including the 5G core and the 5G NR. Here are the main features defined in 33.501.

## Increased Home Control

Home control is used for authentication of the device location when the device is roaming. It allows the home network to verify the device is actually in the serving network when the home network receives a request from a visited network.

This was added in answer to the vulnerabilities found in 3G and 4G networks where networks could be spoofed and send false signaling messages to the home network in an effort to request the IMSI and location of a device. This information could then be used to intercept voice calls and text messages.

## Unified Authentication Framework

In 5G networks, authentication will be access agnostic. The same authentication methods are used for both 3GPP and non-3GPP access networks (5G radio access and Wi-Fi access, for example).

Native support of EAP is key. This allows for new plug-in authentication methods to be added in the future without impacting the serving networks.

## Security Anchor Function (SEAF) and Authentication Framework

5G introduces the concept of an anchor key, with the new function of the SEAF. The SEAF is co-located with the AMF in release 15 of the 5G network specifications. The SEAF allows for the re-authentication of the device when it moves between different access networks, or even serving networks without having to run the full authentication method (e.g., Authentication and Key Agreement (AKA) authentication). This reduces the signaling load on the home network HSS during various mobility services.

The purpose of the primary AKA procedures is to enable mutual authentication between the user equipment and the network that provides keying material reuse between the user equipment and the serving network in subsequent security procedures. The keying material generated by the primary AKA procedure results in an anchor key called the KSEAF provided by the Authentication Server Function (AUSF) of the home network to the SEAF of the serving network.

Keys for more than one security context can be derived from the anchor key without the need of a new authentication run. A concrete example of this is that an authentication run over a 3GPP access network can also provide keys to establish security between the user equipment and a non-3GPP inter-working function used in untrusted non-3GPP access.

The user equipment and the serving network support both EAP-AKA’ and 5G AKA authentication. This is an improvement from previous generations of wireless where different encryption schemes were used depending on the access. In 5G, these two methods are used regardless of the access type, and are the only methods supported.

## Serving Network Authentication

Binding authentication keys to a serving network prevents network spoofing. The serving network thereby authenticates itself to the user equipment. The serving network authentication is used for access technologies (i.e., 5G networks and Wi-Fi access).

## Mitigation of Bidding Down Attacks

5G includes requirements that prevent an attacker from attempting a bidding down attack that makes the user equipment and the network entities respectively believe that the other side does not support a security feature, even when both sides in fact support that security feature.

This prevents fake base stations (also known as IMSI catchers) from forcing devices to 2G access so the IMSI and other data can be more easily captured.

## Subscriber Identifier Privacy

Each subscriber in the 5G system is assigned a globally unique Subscriber Permanent Identifier (SUPI). The SUPI is provisioned in the UDM/UDR and is only used within the 5G network. The SUPI incorporates the IMSI as part of its value, allowing the IMSI to be extracted and used when interworking with 3G/4G networks. The SUPI also incorporates the MCC/MNC for identifying the home network when used in 5G.

The SUPI replaces the IMSI used in previous networks, but the SUPI is never disclosed over the air in the clear when a mobile device is establishing a connection.

In place of disclosing the SUPI, a Subscription Concealed Identifier (SUCI) is used until the device (and network) is authenticated. Only then is the SUPI disclosed by the home network to the serving network but never sent “in the clear” over the air. The MCC/MNC portion of the SUPI is not protected, allowing other networks to quickly determine the home network for the SUPI.



Figure ‑: The user equipment sends a SUCI when the AMF returns an identifier request message in response to a registration or re-registration. SOURCE: 3GPP TS 33.501

The user equipment generates the SUCI when it sends a registration request to the network (if it does not already have a 5G globally unique temporary identifier [GUTI] assigned). The SUCI can also be sent if the network responds to a registration (or re-registration) request with an identity request message.

This procedure has been defined to prevent IMSI catchers (also known as False Base Stations, or Stingrays) from being able to retrieve the subscriber identity by simply attaching to a device. In addition to the SUPI and SUCI, a Permanent Equipment Identifier (PEI) is assigned to the user equipment. This is analogous to the International Mobile Equipment Identifier (IMEI) used in previous generations of wireless networks.

The Subscription Identifier De-concealing Function (SIDF) is responsible for de-concealing the SUPI from the SUCI. SIDF uses the private key part of the privacy related home network public/private key pair that is securely stored in the home operator's network. The de-concealment shall take place at the UDM. Access rights to the SIDF ensure that only a network element of the home network can send a request to the SIDF.

When roaming, the AMF is responsible for issuing a temporary identifier to roamers. The GUTI is allocated to the user equipment on both 5G access and non-3GPP access (i.e., Wi-Fi). The AMF is responsible for managing the assignment of the GUTI and can re-assign the GUTI on new transactions.

## Overview of Security Architecture

The security architecture defined in 33.501 defines security for the following:

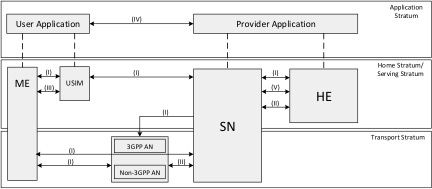


Figure ‑: The 3GPP 5G Security Architecture SOURCE: 3GPP

**Network access security (I):** the set of security features that enable a user equipment to authenticate and access services via the network securely, including the 3GPP access and non-3GPP access, and in particular, to protect against attacks on the (radio) interfaces. In addition, it includes the security context delivery from the serving network to user equipment for the access security.

**Network domain security (II):** the set of security features that enable network nodes to securely exchange signaling data and user plane data.

**User domain security (III):** the set of security features that secure the user access to mobile equipment.

**Application domain security (IV):** the set of security features that enable applications in the user domain and in the provider domain to exchange messages securely.

**SBA domain security (V):** the set of security features about the SBA[[9]](#footnote-9) security including the network element registration, discovery, and authorization security aspects, and also the protection for the service-based interfaces.

**Visibility and configurability of security (VI):** the set of features that enable the user to be informed whether a security feature is in operation or not.

### Security Edge Protection Proxy (SEPP)

To protect messages that are sent over the N32 interface (between two 5G networks), the SEPP is defined as the entity sitting at the perimeter of the 5G network. The SEPP:

* Receives all service layer messages from the network function and protects them before sending them out of the network on the N32 interface and
* Receives all messages on the N32 interface and forwards them to the appropriate network function after verifying security, where present.

The SEPP implements application layer security for all the information exchanged between two network functions across two different PLMNs.

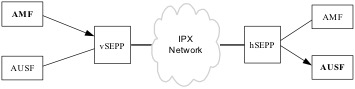


Figure ‑: The role of the SEPP in the security architecture. SOURCE: 3GPP

Protection for the HTTP message payload, sensitive information contained in the HTTP message header, and the Request URI is provided by the SEPP. However, not all information in the payload gets the same level of protection. Some information may require end-to-end encryption, while other information may only require integrity protection end-to-end, while still allowing modification of the message by intermediary Internetwork Packet Exchange (IPX) providers.

The SEPP provides confidentiality of sensitive information as it is passed between two networks. This prevents intermediary networks (such as IPX) from being able to see this sensitive information (such as authentication vectors).

If there are parameters in the data that need modification by the IPX (such as parameters modified by mediation for interoperability), both networks must agree on what changes can be made and those changes must be integrity protected. When modifications are made by an IPX, the receiving network verifies the changes. The SEPP can detect any unauthorized message modifications received by an IPX.

The SEPP also provides traditional gateway functions, such as topology hiding, protection from malformed packets, protection against signaling storms, and validation of the message source.

## New Security Requirements on the user equipment

User equipment changes have been defined for 5G as well. The subscribers’ credentials are protected using a tamper resistant hardware component, to prevent changing the identity of the device or accessing sensitive credentials.

Another capability in 5G for the user equipment is the ability to manage the radio technology used on the user equipment.

The user equipment allows the subscriber to disable and re-enable one or more of the radio technologies supported by the device, regardless of the network. These could be 4G/LTE, GSM/EDGE, WCDMA, or 5G NR.

The home network also can disable and re-enable radio technologies on the device, such as restricting the use of GSM/EDGE when 5G NR is available. This must be done using a secure connection between the user equipment and the network. These settings are remembered even when a user equipment powers down or is rebooted. If a prioritized service (such as emergency services, MPS, or mission critical services) is needed, and the user equipment is unable to connect to a network using the radio technologies that have been allowed, the prioritized service can override the user and network and select the radio technology supported an accessible network.

# The Internet of Things (IoT) in 5G

The security of the IoT is manifested in a broad ecosystem that is comprised of many participants including network operators, manufacturers, software developers, and service providers. In the 5G environment, this ecosystem will continue to expand to include vehicles, industrial control systems, and sensors driving intelligent municipalities (smart cities) to name but a few.

## Previous Work Efforts

There has been a substantial focus on identifying threats, mitigating actions and best practices to address the cyber security of the Internet of Things. Numerous reports have been generated as a result of the work of these public/private partnerships. Since much of this content will be directly applicable in the 5G environment, some of these reports are referenced below.

### CSRIC IV, WG5 Remediation of Server-Based Distributed Denial of Service (DDoS) Attacks

Working Group 5 provided recommendations that communications providers can take to mitigate the incidence and impact of DDoS attacks from data centers and hosting providers, particularly those targeting the information systems of critical infrastructure sectors. The recommendations are mainly in the form of server-based DDoS mitigation best practices. In addition, several actionable recommendations were included to further the work to prevent, detect, and mitigate server-based DDoS attacks.

### CSRIC III, WG7 Botnet Remediation and the Anti-Botnet Code of Conduct

The CSRIC III tasked Working Group 7, Botnet Remediation, with proposing a set of agreed upon voluntary practices that would constitute the framework for an opt-in implementation model for Internet Service Providers (ISPs) to follow to mitigate the botnet threat. In response, the U.S. Anti-Bot Code of Conduct for ISPs was developed to address the threat of bots and botnets in residential broadband networks through voluntary participation. It was determined in developing the code that constituents of the entire Internet ecosystem have important roles to play in addressing the botnet threat and that ISPs depend on support from the other parts in the ecosystem.

### CSRIC II, WG8 ISP Network Protection Practices

Working Group 8 of CSRIC II addressed the area of ISP Network Protection, with a focus on addressing “bots” and “botnets.” Botnets are formed by maliciously infecting end-user computers and other devices with malware through a variety of means, and surreptitiously controlling the devices remotely to transmit onto the Internet spam and other attacks (targeting both end users and the network itself). The Working Group examined potentially relevant existing Best Practices (BPs), and in consultation with industry and other experts in the field, identified additional best practices to address protection for end-users as well as the network. The best practices identified were primarily for use by ISPs that provide service to consumer end-users on residential broadband networks but may apply to other end-users and networks as well.

### Communications Sector Coordinating Council White Paper

The Communications Sector Coordinating Council (CSCC) published a technical white paper specifically to inform the process undertaken by the Government in response to the Executive Order 13800 by describing the shared responsibilities of key participants in the Internet ecosystem for mitigating the threats posed by botnets.

The Internet ecosystem has been working collaboratively to neutralize the threats from bots and botnets for years. In this paper, the CSCC identified a number of challenges of mitigating botnets, and opportunities for increased collaboration and cooperation among members of the Internet ecosystem to address the problem.

### Report to the President on Enhancing the Resilience of the Internet and Communications Ecosystem Against Botnets and Other Automated, Distributed Threats

In recognition of the growing global threat represented by malicious botnets, President Trump signed Executive Order 13800, *Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure,[[10]](#footnote-10)* tasking the Department of Commerce (DoC) and the Department of Homeland Security (DHS) to lead an open and transparent process to identify ways to improve the resilience of the internet and communications ecosystem and reduce the threats perpetuated by bots and botnets. The final report was transmitted by the Secretary of Commerce and the Secretary of Homeland Security on May 22, 2018. This final report was the result of a joint effort involving three approaches – hosting workshops, publishing requests for comment, and initiating an inquiry to the President’s National Security Telecommunications Advisory Committee.[[11]](#footnote-11) The resulting report included recommendations and identified challenges to reduce the threat from automated, distributed attacks.

## IoT Service Enablement in 5G

It is not the intent of this report to repeat the work stated above. Rather, this section of the WG3 report focuses on the essential interconnectivity between the “Things” and the “Internet” that will be enabled as a result of the realization of a 5G network. Earlier sections of this report and prior work from WG3 address many of the new impactful technologies and capabilities enabled by 5G such as the evolution of the EPC, NFV/SDN capabilities and the Diameter signaling protocol. The specific focus of this section of the report is the discussion of vulnerabilities and mitigation of vulnerabilities in the software-based SOA and the application programming interface (API) that will be used for the development and deployment of the vast new pool of services that will enable the evolution to the “Internet of Everything”. In addition to the discussion on use cases and service enablement, this section also describes the unique interconnectivity enabled by Narrowband IoT and the introduction of specific platforms and interfaces to securely enable this use-specific communications capability within the 5G network.

## Examples of IoT applications[[12]](#footnote-12)

IoT encompasses so many different ‘things’ in our lives today. According to several analysts, it is expected that IoT will reach 50 billion devices by 2020. When one thinks about all of the different uses for connected things, it is not hard to fathom.

There are connected things that are industrial, some that are medical, and some that are consumer products. All of them have one thing in common; they represent another attack vector that can be used against critical infrastructure if not managed and secured.

### Industrial Process Automation

Industrial processes, localized or geographically distributed, are increasingly automated to ensure quality, consistency, and cost-effective production of goods or services. Automation systems for these processes broadly consist of instrumentation, control, human interface, and communication subsystems. These processes are in general spread over a large area and utilize many sensors and actuators to monitor and control complex processes that are connected with each other. Connectivity is required for these sensors and actuators both indoors and outdoors with high availability and reliability to ensure seamless production and the ability to adapt processes in real-time for maximum flexibility.

### Autonomous Driving

Fully autonomous driving involves the capability of a vehicle to sense its environment and navigate without human input under all scenarios and conditions. This is an evolving development area and it is anticipated that by 2020, 98% of cars sold will be connected to the Internet.[[13]](#footnote-13) Autonomous cars use a combination of technologies to detect their surroundings including wireless communication technologies, laser and radar sensing, GPS, odometers, computer vision, and advanced control systems. All this data is analyzed, processed with artificial intelligence and deep learning computer systems to distinguish between different cars on the road and identify appropriate navigation paths given obstacles and considering the rules of the road. 5G technologies are anticipated to enable these cooperative automatic driving use cases in an enhanced fashion where sensor information will be exchanged in real time between thousands of cars connected in the same area. The 5G features are expected to provide communications with increased coverage, reliability and performance levels with higher orders of magnitude and much lower latency compared to existing technologies today.

### Smart Grids

Improved communication technology will enable enhanced monitoring, better management, and greater control of energy generation and distribution networks leading to increased availability and resilience. Energy generation and distribution are shifting towards decentralized and smaller power generation sources with more variable power delivery that will require tighter integration and secure communication networks. Also, new storage solutions are used as part of the delivery process with new battery materials offering higher energy storage capacities. With these new energy generation and storage scenarios emerging, there is an expansion of usage patterns and end devices such as the proliferation of electric vehicles and IoT. Low latency and ultra-reliable 5G is but one technology that is anticipated to address the communication requirements for Smart Grids.

### Remote Surgery

The capability of a surgeon to remotely operate a surgical robot to perform surgery on a patient will allow more uniform access to talented surgeons and better utilize their skills.

The transfer of high-resolution images and video to the surgeon requires large bandwidth on the uplink. In event of an emergency, high availability of the necessary robots and surgeons certified for their use will also be required. QoS guarantees for low latency and reliability requirements are critical.

### Media and Entertainment

The media and entertainment industry seeks to improve the user experience and enable access to an expanding universe of content anytime and anywhere. This vertical opportunity focuses on different types of multi-media services that include regular live/linear media, on-demand content, user-generated content and gaming. Due to consumer demand, media use needs to meet both stationary and mobile end users. The end users are also consuming media on an increasing variety of devices that include TVs, smartphones, tablets, wearables and other devices.

This vertical increasingly requires higher data rates to provide high resolution multimedia content to an increasing number of simultaneous and connected users with high QoS requirements. With more devices capable of recording and capturing our daily experiences, there has been a dramatic increase in user-generated content. The popularity of social media sharing and the growing size and scale of platforms like Facebook, Instagram, and Snapchat, among others, are also driving requirements for increased uplink data rates.

## Taxonomy of Use Cases

5G is well suited for the transformation of the vertical industries outlined in the preceding section, supporting a wide variety of applications and use cases with high variability in key performance attributes such as mobility, data rate, scale, latency and reliability. For example, mobility could range from an application like fixed wireless service to connected vehicles that may be moving at speeds of 80 miles per hour. Data rates could vary across a similar range from bits per second for some IoT devices to gigabits per second for virtual reality. The ultra-low latency needed to enable real-time applications like industrial automation is different from smart home applications that may be more delay tolerant. The rapidly increasing number of devices will necessitate the ability to rapidly scale services. Reliability is critical for remote surgery and healthcare monitoring but maybe less so for some remote sensors and meters in smart cities.

In the previous section, a few use cases for different verticals were described and it is notable that even use cases within the same vertical can have distinct KPIs. As an example, an automated product line use case in the industrial automation vertical requires low latency and highly reliable communication. These are different KPIs when compared to the use case on inventory and supply chain optimization in the same vertical, which requires many sensors and the latency requirements are much less strict. As a result, use cases can be categorized based on their performance attributes.

1. *Enhanced Mobile Broadband (eMBB):* These use cases generally have requirements for higher data rates and better coverage
2. *massive Machine Type Communications (mMTC):* These use cases generally have requirements to support a very large number of devices in a small area, therefore, they support a large device density
3. *Critical communications:* These use case have strict requirements on latency and reliability, and are also referred to as Ultra Reliable and Low Latency Communications (URLLC)

With such a large variation in performance attributes, it is also useful to consider these different use cases in terms of their types of interaction: between people, between machines, or between people and machines.

Combining these types of interaction and grouping the use cases by the primary categories that 5G will impact-- extreme mobile broadband, massive scale communication and ultra-reliable low latency service-- creates a powerful alternative visualization. It enables a vision of the way certain use cases will span across multiple types of interaction and various performance requirements. Figure 5-1 shows this new taxonomy for some 5G use cases.

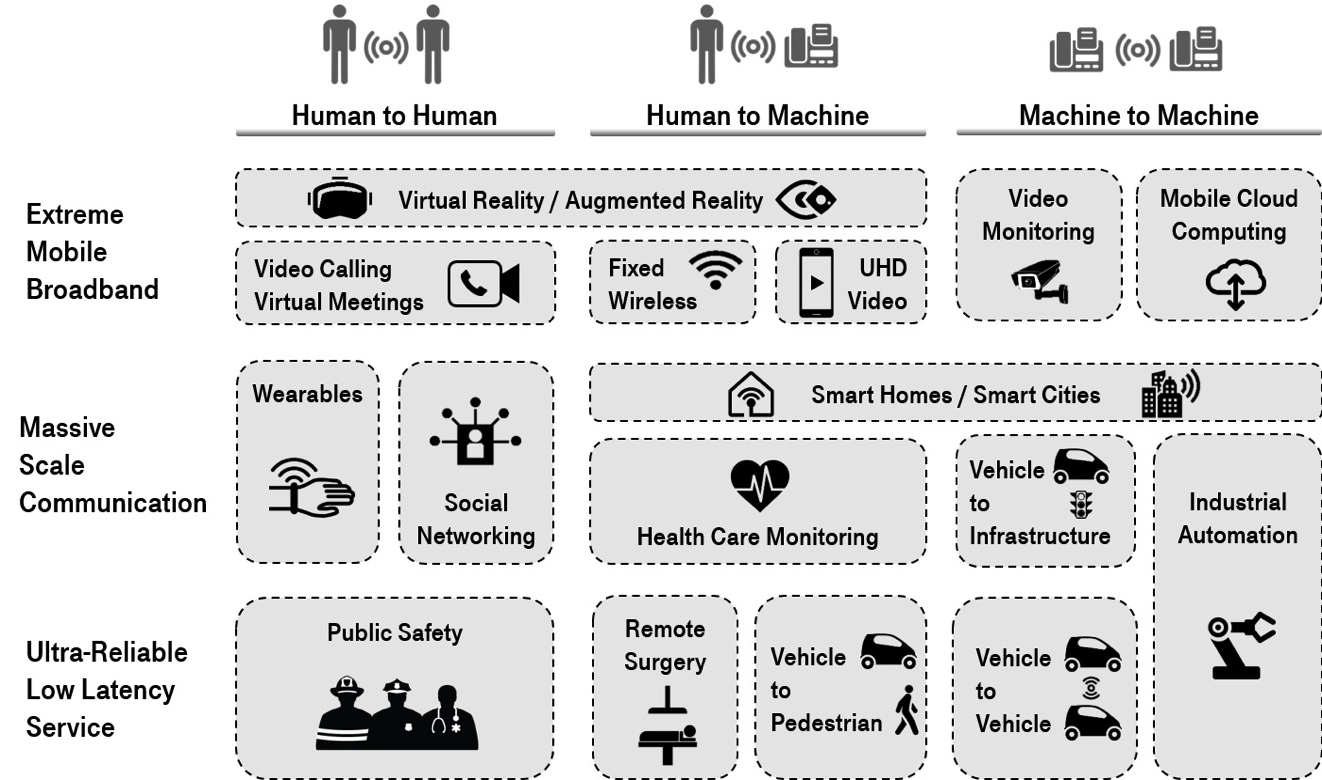


Figure ‑: 5G use cases grouped by the type of interaction and the range of performance requirements. SOURCE: page 19 of 5G Americas Services & Use Cases.

The use cases associated with the mMTC depicted, as Massive Scale Communication in the figure above typically will leverage the emerging Low Power Wide Area (LPWA) needs for low cost devices, extended coverage, and long battery life. The use cases are expected to make up a large part of the new types of services that 5G systems will address by connecting the massive number of devices such as sensors, actuators, cameras, and wearables.

This family of use cases is expected to be pervasive in “smart” cities, as buildings and industries will evolve that leverage these devices providing metering, lighting management in buildings and cities, environmental monitoring (pollution, temperature, noise, etc.) and traffic control, among many other applications.

These services are expected to require the ability to support a high density of devices with different characteristics in a common communication framework. The mMTC use case category includes applications used in a wide spectrum of industries across society, including both human-to-machine interaction and machine-to-machine interaction, as shown in the following figure:

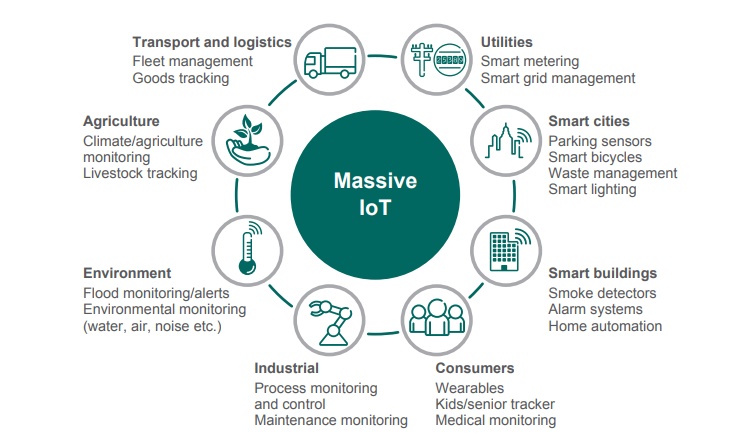


Figure ‑: Several use cases in mMTC category enabled by 5G technologies. SOURCE: 5G Americas Services & Use Cases, page 22.

## GSMA IoT Model

The GSMA has published a set of IoT Security Guidelines as non-binding reference documents for the benefit of service providers looking to develop and deploy new IoT services. These guidelines were developed to help ensure that new IoT services coming to market are made more secure. The industry partnership that developed these guidelines included network operators, application/service providers, and device equipment manufacturers. The structure of the resulting document set, the GSMA IoT Security Guideline Document Set, is shown below:[[14]](#footnote-14)

CLP.17 GSMA IoT Security Assessment Checklist

**CLP.13**

IoT Security Guidelines for IoT Endpoint Ecosystem

**CLP.12**

IoT Security Guidelines for IoT Service Ecosystem

**CLP.11**

IoT Security Guidelines Overview

Document

**CLP.14**

IoT Security Guidelines for Network Operators

Figure ‑: GSMA IoT Security Guidelines Document Structure. SOURCE: GSMA

The figure below shows the standard IoT model used throughout the GSMA Reference documents. This diagram defines the primary components that are required when deploying a production-ready IoT service or technology. Communications networks are inherent to IoT and, for the purposes of this model, provide the connection between the endpoint ecosystem and service ecosystem.

****

Figure ‑: Example IoT Model. SOURCE: GSMA

Specific network security guideline recommendations for network operators can be found in the GSMA’s “IoT Security Guidelines for Network Operators”.

### Service Ecosystem

The service ecosystem represents the set of services, platforms, protocols, and other technologies required to provide capabilities and collect data from endpoints deployed in the field. This ecosystem typically gathers data from endpoints and stores them within its server environment. This data, often in the form of metrics, parameters or commands, can also be handed off to authorized third parties via an API (e.g. oneM2M) originating at the service infrastructure, which is commonly how IoT service providers monetize the service.

### Endpoint Ecosystem

The endpoint ecosystem consists of the IoT devices and gateways that connect the physical world to the digital world. Endpoints gather metrics from the physical environment around them and push that data in different formats using attached communications networks to the service ecosystem, often receiving instructions or actions in response.

## IoT Connectivity in 5G

One of the characteristics of IoT communications is the broad spectrum of capabilities already discussed earlier in this report. The number of these devices is anticipated to become quite large, up to several magnitudes of order compared to traditional device connections. Using existing LTE technology would lead to a network overload, because despite the small amount of user data transmitted by many of these devices, the amount of network signaling is essentially unchanged. In response, 3GPP developed specifications for a new type of communications that would leverage the core network while minimizing both radio and signaling impact to the network.

The NB-IoT specifications also embraced device benefits include lower power consumption requirements and network-based authentication and authorization controls. The result is that there are two methods in the 5G IoT to send data to an IoT application through the EPC and later the 5G core, the user plane and the control plane.

In the control plane method, data are transferred from the RAN to the MME. From there, they may either be transferred via the serving gateway (SGW) to the packet data network gateway (PGW), or to the SCEF (only possible for non-IP data packets). From these nodes, they are finally forwarded to the service capability server (SCS) that connects the IoT application servers to the network operator core network.

With the user plane method, data is transferred in the same way as the conventional data traffic, i.e. over radio bearers via the SGW and the PGW to the application server. Thus, it creates some overhead on building up the connection, however it facilitates a sequence of data packets to be sent. This path supports both, Internet Protocol (IP) and non-IP data delivery.

This same capability will exist in 5G networks. The SCEF evolves to the Network Exposure Function (NEF) and is still responsible for the management of low power IoT (NB-IoT). There are other new functions introduced in 5G that will be involved in the transfer of data between IoT devices and their associated application functions, but management of the device remains the role of the NEF.

### How NB-IoT Flows

In June 2016, 3GPP ratified the NB-IoT standard as a low-power solution that will work with existing wireless infrastructure. NB-IoT is optimized for machine type traffic. It is kept as simple as possible in order to reduce device costs and to minimize battery consumption. In addition, it is also adapted to work in difficult radio conditions, which is a frequent operational area for certain machine type communication devices. Although NB-IoT is an independent radio interface, it is tightly connected with the LTE core network, which also shows up in its integration in the current LTE specifications.

In this solution, there is no data radio bearer set-up; data packets are sent on the signaling radio bearer instead. Consequently, this solution is most appropriate for the transmission of infrequent and small data packets. The SCEF is used for the delivery of non-IP data over control plane and provides an abstract interface for the network services (authentication and authorization, discovery and access network capabilities).

In 5G networks, the SCEF evolves to the NEF and provides a means to securely expose the services and capabilities provided by 3GPP network interfaces. The NEF provides access to network capabilities through homogenous network application programming interfaces (e.g. network API) defined by OMA, GSMA, and other standardization bodies. The NEF abstracts the services from the underlying 3GPP network interfaces and protocols.[[15]](#footnote-15)



RAN

MME

SCEF



SGW/PGW

Non IP Data

IP Data

IP and Non IP Data



SCS



AS

Figure ‑: Role of the SCEF and SCS in 4G networks. SOURCE: T-Mobile

The SCS is an entity that connects the various IoT application servers to the mobile network operator’s network to enable them to communicate through specific 3GPP defined services. The SCS can be controlled by either the mobile network operator or by the IoT service provider and can connect to one or more IoT application servers.

## Risks of IoT in 5G

The risks associated with IoT are many. The propagation of malware between devices continues to be a major issue for IoT, and the distribution of botnets will only continue as 5G networks are deployed.

There are many different ways for protecting the network from misbehaving IoT devices, as well as protecting the IoT network itself.

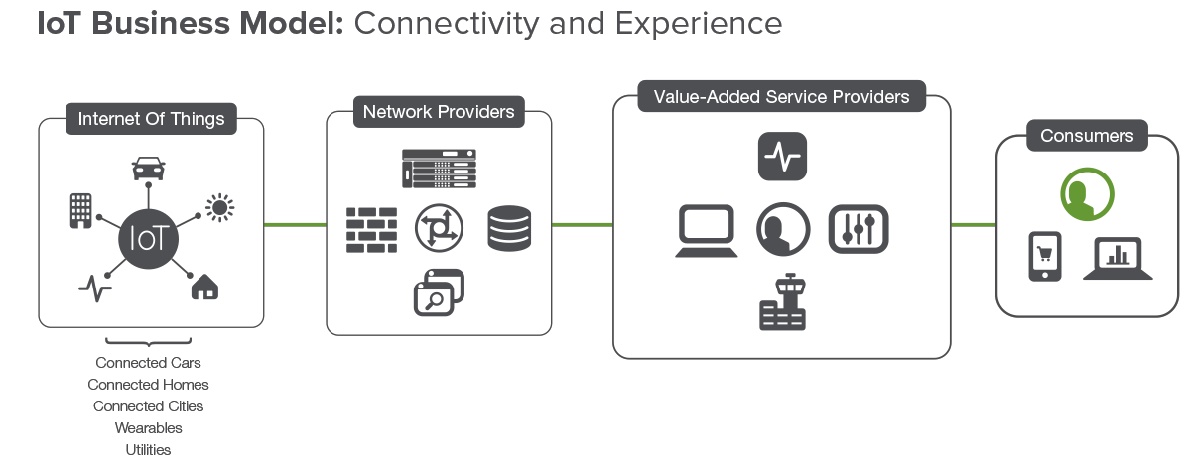


Figure ‑: The IoT business model. SOURCE: F5 Internet of Things: Security and Business Impacts on Service Providers.

### IoT Service Assets That Network Operators Can Protect

The security features that need to be implemented to adequately protect IoT service assets are specific to each service. Therefore, it remains the responsibility of the IoT service provider to use proper risk and privacy impact assessment processes to derive their specific security needs. Network operators and IoT service providers often share similar security requirements to protect their assets, therefore it makes sense for them to use common security solutions rather than implementing duplicate (and potentially redundant) security infrastructures. Moreover, in many cases the network operator will also be the IoT service provider. The security services provided by network operators can provide a critical role in securing the assets used to provide an IoT service. These assets can include:

* IoT service data being sent between an IoT endpoint device and the IoT service platform – this includes both primary privacy-sensitive data (e.g. end user related data) and commercially exploitable data (e.g. such as actuator control data) which may also have some secondary privacy impact.
* The security assets (IMSI, keysets etc.) and network configuration settings (APN, timer values etc.) used within endpoint devices (including gateway devices).
* IoT service provider’s business-sensitive information, including brand reputation, customer/user data under company responsibility, strategic information, financial data and health records, etc.
* An IoT service provider’s business infrastructures, service platforms, corporate networks and other private network elements.
* Public (i.e. shared) data center infrastructures provided by the network operator that are used by the IoT service. This can include public services, hosted capabilities, virtualization infrastructures, cloud facilities, etc.
* Communications network infrastructure, including RANs, core network, backbone networks, basic service functions (DNS, BGP, etc.), access to and aggregation of fixed and cellular networks, etc.

New IT technologies, like virtualization and SDN/NFV, are seen as a way to make 5G networks more nimble and efficient, yet less costly.

Security cannot be built for 5G services unless the network infrastructure is robust. In legacy networks, security of network functions (NFs) relies largely on how well their physical entities could be isolated from each other. However, in 5G, the isolation will work differently for virtual network functions on cloud-based infrastructure.

SDN is proved to be of help in improving transmission efficiency and resource configuration. On the other hand, it is important to consider in the 5G security design that it could be managed in terms of the isolation for network nodes such as control nodes and forwarding nodes, and the secure and correct enforcement of the SDN flow table.

Based on network virtualization technology, a network can have different virtual network slices. Each virtual network slice could accommodate a particular service requirement and that may require differentiated security capabilities. 5G security design may need to consider issues of how to isolate, deploy, and manage virtual network slices securely.

IoT devices have many choices in the way they access networks. For instance, they may connect to networks directly, or via a gateway or controller, or in relay fashion.

Comparing to mobile handsets, security management of IoT devices in 5G may be efficient and lightweight in order to establish trust relationships between devices and networks.

While the data transported will be subject to privacy protections, another consideration will be the use of automated control mechanisms based on specific service type sensing that may also involve access to personal identity information. Personally Identifiable Information such as location, or even sensitive Private Health Information, may be utilized in the service type sensing process.

This injects an additional layer of privacy protection requirements in 5G for IoT to help secure data in transport and protecting information used to enable service type sensing.

End-to-end security design caters to different vertical industries. In that case, the design of security protection needs to consider how to fulfill security requirements of the specific industry segment.

Network operators will need the tools to tailor security protection to the requirements of the application. For example, robotic surgery and autonomous vehicles will require significantly more rigorous security measures than embedded street sensors.

To provide better support and rapid response to the vertical industry requirement, end-to-end security capabilities could be rapidly aligned with business changes. In that case, it would require flexible and highly efficient end-to-end security deployment and adaptation.

5G will see massive growth in personal privacy data used by application services including device identifiers, user IDs, and user preference. Privacy protections should be built end-to-end, reducing vulnerabilities to privacy leaks.

The telecom industry is seeking to better serve vertical industries. The industry has done well in protecting user privacy, and users have developed a level of trust with the security strength of the communication systems. The industry is poised to continue to extend the user trust by opening up security delivery capabilities as a service to individual users and vertical industries.

As established IT technologies (e.g. NFV and SDN) are put into use, a vast array of system-level protection is made available to defend against DDoS and other active attacks that may increase.

Both software and hardware infrastructures run in multi-vendor environments. In order to mitigate unauthorized access to network resources, stringent identity management is needed.

Integrity and confidentiality protection are provided throughout data transmission to prevent data from being intercepted or re-routed to unauthorized destinations.

The evolving trust model for 5G IoT services will involve three distinct constituents: Users, Networks operators, and Service providers.

Three authentication models will likely co-exist in a 5G IoT environment to support the diversity of platforms and services.

* Network only authentication where the network operators offer service providers authentication services to facilitate single sign-on flexibility for users.
* Service provider authentication (vertical industries) that relies on the service provider to provide authentication that exempts the device authentication process and reduces network costs.
* Dual authentication (network operator for access, service provider for services) represents the legacy model where network operators handle network access authentication and service providers manage service access.

Taking into account that many of the IoT devices will not have much processing power, edge authentication will be a significant challenge; especially when interacting with automated processing devices that do not have human interaction (how do you implement “what you know” and “what you have” type of authentication?).

Manufacturers of many types of IoT devices are sometimes small companies with few resources at their disposal. In an attempt to get to market as quickly (and cheaply) as possible, security is not built-in to the device. This is especially true with the many consumer products on the market today.

Operators should also anticipate that IoT devices will not be upgradeable, and version control by the vendor non-existent. Life cycle management is not a concept for many of these small manufacturers.

It is also important to recognize that many IoT devices – many of which have not, as stated, been designed with security in mind – will reside in networks for much longer periods of time relative to the devices that have historically been attached to networks. The model, with which many network operators are familiar, in which devices are replaced every twelve to twenty-four months, will no longer apply to many of the IoT devices in the future.

There are two categories of IoT communications; massive data volumes from low cost devices and critical data from time sensitive or high availability devices such as vehicle-to-vehicle (V2V), traffic safety/control, and medical devices. These impact security considerations differently in that the first is a scale challenge and the second is a QoS/performance challenge. There is also a two-pronged issue for network operators: (1) many IoT devices will not have the capability to protect themselves due to cost and life-cycle issues, and (2) the proliferation of devices will be on a scale not familiar in legacy networks (connected devices measured in the billions).

IoT devices critical to lives require real time availability. Device-to-device and network-to-device DoS are a real risk for IoT devices like medical devices, and cars. Regardless, the security model changes from end device authentication and protection to “in flight” protection and verification of the data itself.

Historically, the network endpoints were the focus of protection (think firewalls, anti-virus, etc.) and the interconnection of network elements internal was through a “trusted” network; this model needs to change.

The actual transport has to guarantee the accuracy of data transmitted, verify the source of the data transmitted, and confirm delivery of the data transmitted; all of this while ensuring that in transit processing does not impact latency or propagation

This translates into an integrated vision where network properties of QoS, availability, and security are considered together in order to support a target resilience level.

## Threat Assessment for IoT

The most fundamental security mechanisms provided by a communication network are:

* Identification and authentication of the entities involved in the IoT service (i.e. gateways, endpoint devices, home network, roaming networks, service platforms).
* Access control to the different entities that need to be connected to create the IoT service.
* Data protection in order to guarantee the security (confidentiality, integrity, availability, authenticity) and privacy of the information carried by the network for the IoT service.
* Processes and mechanisms to guarantee availability of network resources and protect them against attack (for example by deploying appropriate firewall, intrusion prevention and data filtering technologies)

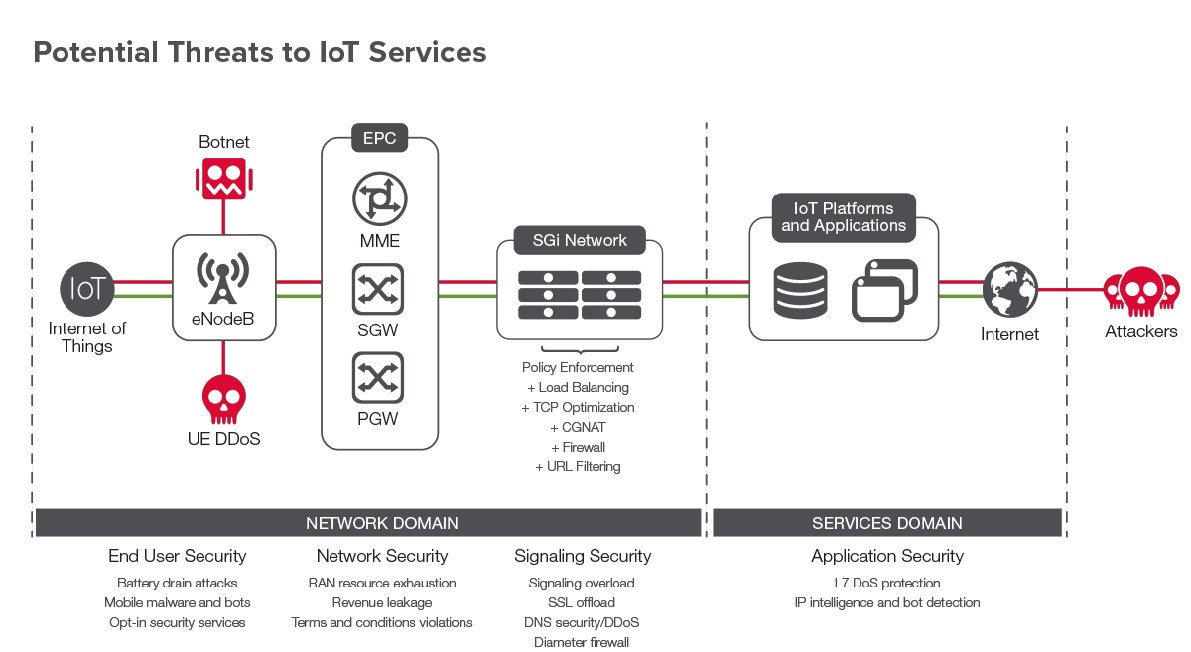


Figure ‑: Potential threats to IoT services. SOURCE: F5 Internet of Things: Security and Business Impacts on Service Providers.

In a February 2016 survey and report on The Future of Mobile Service Delivery, Jim Hodges, senior analyst for Heavy Reading, revealed global service providers' priorities and concerns and explored how they can address security while preparing for a 5G/IoT world. This report cited that the top three security concerns reported by service providers in decreasing order of importance are:

* Denial of availability (this includes DDoS and botnet driven traffic)
* System integrity violations
* Identity spoofing[[16]](#footnote-16)

### Emerging Security Threat Environment in the 5G-Based IoT Service Oriented Architecture

IoT middleware is a software layer or set of sub layers that decompose the complexity of the underlying systems into components that abstract from the specifics of the underlying hardware, radio, and other technologies. This functionality is then presented through software as standard services typically using a web services application programming interface (API). The middleware’s ability to hide the details of different technologies is fundamental to exempt the programmer from issues that are not directly pertinent to his/her focus, which is the development of speciﬁc applications enabled by IoT systems infrastructure.

Many of the system architectures proposed comply with the SOA approach. The adoption of SOA principles allows the decomposition of complex systems into applications consisting of a system of simpler and well deﬁned components. In a SOA architecture, each system offers its functionality as standard services. Moreover, the SOA architecture supports open and standardized communication through all layers of web services.

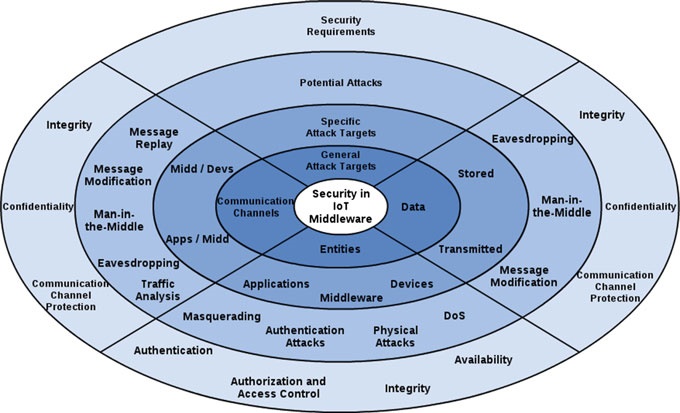


Figure ‑: Security Taxonomy. SOURCE: ResearchGate White Paper, Figure 3, page 406

The figure above presents a security taxonomy for SOA-based IoT middleware which identiﬁes the most attractive targets for future attackers in the upcoming 5G communications systems and the security requirements for these systems. According to the taxonomy, the attacks can occur in entities, data and communication channels. Attacks against entities are related to unauthorized access in applications, middleware or devices. With the expected growing of the devices combined with the increased data transmission capabilities of 5G networks, the wide adoption of open operating systems and the fact that the devices will support a large variety of connectivity options are factors that render these entities a prime target for attacks. Devices and mobiles technologies will suﬀer with DoS attacks via messages, malware, etc. An invaded device can compromise an entire network where it is inserted. Potential attacks in entities are masquerading, authentication attacks, physical attacks and DoS attacks.

Since IoT is becoming omnipresent, privacy has become a concern. It is extremely important to disclose user’s data only to authorized parties. In this sense, data attacks can happen in two ways: when data are changed or spied during the transmission between entities, and/or when the stored data are illegally modiﬁed in the data repository. Potential attacks in data are message modiﬁcation, eavesdropping and man-in-the-middle. These attacks happen on the communication channel; however, retrieving data is the main goal.

The fact that 5G mobile systems will support many different access networks leads them to inherit all the security issues of the underlying access networks that they will support. 5G mobile systems will be vulnerable to communication channel attacks that are common over the Internet. They happen in the communication between system entities. An IoT middleware basically has two communication channels, one with applications and another with devices. Both channels can be exploited by attacks. External networks can be the targets of DDoS attacks in 5G communications systems, where mobile botnets generate a high volume of traffic and transmit it to the target over the core network. Another potential attack is message insertion, an attack that occurs in LTE networks when an attacker injects control protocol data units into the system to achieve DoS attacks against a device. Other potential attacks in communication channels are eavesdropping, man-in-the-middle, message modiﬁcation, and message replay and traffic analysis.

### REST API Threat Environment[[17]](#footnote-17) [[18]](#footnote-18)

The use of APIs to enable applications to interact across single and multiple corporate infrastructures is an ever more widespread activity. Indeed, it is now common to refer to an "API economy" in which companies are finding new and innovative ways of monetizing their software assets by exposing APIs for developers to harness the power of their features and functions.



Figure ‑: The API economy is a reality today, “API Security: A Disjointed Affair”, SOURCE: Rik Turner, Ovum, 7 April 2016

REST (REpresentational State Transfer) is an architectural style that can be used to communicate with web services. In many ways, REST has a lot in common with protocols such as SOAP; it is used as a communication mechanism between two applications, or between an application and an online service. Many mobile web applications communicate with a RESTful API at the backend to interact with the online service.

The key abstraction of information in REST is a resource. A REST API resource is identified by a URI, usually an HTTP URL. REST components use connectors to perform actions on a resource by using a representation to capture the current or intended state of the resource and transferring that representation. The primary connector types are client and server; secondary connectors include cache, resolver and tunnel. In order to implement flows with REST APIs, resources are typically created, read, updated and deleted. For example, an ecommerce site may offer methods to create an empty shopping cart, to add items to the cart and to check out the cart. Another key feature of REST applications is the use of standard HTTP verbs and error codes in the pursuit of removing unnecessary variation among different services. Secure REST services must only provide Hypertext Transfer Protocol Secure (HTTPS) endpoints. This protects authentication credentials in transit, for example passwords, API keys or JSON Web Tokens. It also allows clients to authenticate the service and guarantees integrity of the transmitted data.

Relaxed security in legacy internal-facing SOA services is ill equipped to face the hostile environment that is the Internet. Even if security was built into the internal services it is often made obsolete by new threats. Furthermore, most enterprise backend systems comprise a dizzying array of incompatible legacy- and acquired technologies. Re-architecting or retro-fitting security in such systems is often not possible, especially given time-to-market pressures.

Irrespective of goals, motivations, and technology or architectural preferences, APIs are just another front end to core backend services. A majority of users may already be reaching an entity through myriad apps exercising APIs, rather than through browsers. However, the backend services and their security requirements remain the same.

The evolution of APIs has ushered new paradigms of interactions over HTTP, which traditional security technologies like SOA gateways, Intrusion Prevention Systems (IPS) and Intrusion Detection Systems (IDS) and web application scanners struggle to manage. As an example, the use of JSON in HTTP requests presents new conduits through which untrusted data can reach backend services or an end user’s browser, where it is consumed. Similarly, would-be attackers can pass user inputs within the URL path (rather than URL query) with REST, breaking legacy security tools.

Another challenge with APIs is service availability. APIs are exercised programmatically and can be chatty. Unanticipated use, verbose applications or abusive partners can wreak havoc upon the API SLAs (Service Level Agreements), or even bring down the backend services, with severe financial implications.

With the growth of REST APIs inside enterprises but also outside their boundaries in their ecosystem, monitoring, protecting and preventing attacks is key and REST API security is of paramount importance. Not only failures in security implementations get API project stakeholders on alert but also regulations like PSD2 have been kick-starting initiatives to standardize security implementations. Questions around counter measures and best practices in API security are now even getting attention from top level management, due to the dramatic impact a security breach might potentially have on the company profitability and reputation.

### Potential Attacks and Security Countermeasures to Protect the 5G IoT Middleware architecture

Table 5-1 depicts the relationship between the security requirements for 5G-based IoT middleware architecture and the attacks. The attacks can be divided between active and passive. Active attacks are those carried out by transmitting or replaying traffic, while passive ones are only based on listening to traffic.

Active attacks are:

* ***Man-in-the-Middle*:** Attacker intercepts the path of communications between two legitimate parties, in an attempt to obtain authentication credentials and data.
* ***Message Modiﬁcation*:** Attacker actively alters a legitimate message by deleting, adding to, changing, or reordering it.
* ***Masquerading*:** Attacker impersonates an authorized user and in an attempt to gain certain unauthorized privileges.
* ***Authentication Attacks*:** Intruders use these attacks to steal legitimate user identities and credentials. Dictionary attacks and brute force attacks are two common attacks in this category.
* ***DoS Attacks*:** Attacks attempt to inhibit or prevent legitimate use of the communication services.
* ***Physical Attack*:** Attacker has physical access to the device and can steal credential information like static keys.
* ***Message Replay*:** Attacker passively spoofs transmission frames and retransmits them, acting as if the attacker is a legitimate user. This attack is also considered an active attack.

Passive attacks are:

* ***Eavesdropping*:** Attacker passively monitors the network communications for capturing communicating data and authentication credentials.
* ***Traffic Analysis*:** Attacker passively monitors transmissions to identify communication patterns and participants.

In order to protect the whole 5G-based IoT middleware architecture from these attacks, some security countermeasures must be developed and deployed in the middleware architecture. The next topics describe a set of security requirements that should be used to ensure protection of the whole architecture:

* ***Authentication*:** It is necessary to establish an authentic connection between two entities in order to exchange data and keys in a reliable manner. In IoT context, mutual authentication is required because IoT data are used in different decision making and actuating processes. Therefore, both entities need to be assured that the service is accessed by authentic parties, and service is offered by an authentic source. Furthermore, strict authentication mechanisms need to be deployed in order to prevent impersonation. Enforcing any authentication mechanism requires to register user identities. Moreover, the resource limitation of IoT objects poses stringent constraints to enable any authentication technique.
* ***Authorization*:** It refers to the means of expressing the access polices that explicitly assign certain permissions to subjects based on previous authentication. The IoT environment needs to provide ﬁne-grained, reusable, dynamic, easy to use policies deﬁning and updating mechanism. Thereby, it is imperative to externalize the policy deﬁnition and enforcement mechanism of IoT services. Authorization is a mandatory requirement for applications and devices since they have different privileges for accessing speciﬁc resources and services in an IoT middleware.
* ***Access Control*:** This is an enforcement mechanism that allows only authorized users access to the resources. The enforcement is usually based on access control decisions. Since IoT is becoming omnipresent, privacy has become a real concern. It is important to disclose user’s data only to authorized parties.
* ***Communication Channel Protection*:** The role of this requirement is to protect the communication channels between applications/devices and middleware. The goal is to protect the data exchanged by entities against attacks during the transmission through the use of security protocols that must ensure the communication channel protection independent of the security requirements used by them.
* ***Conﬁdentiality*:** This requirement can be achieved through cryptography mechanisms. Different existing symmetric and asymmetric cryptography schemes can be leveraged to ensure conﬁdentiality. The selection of a particular cryptography algorithm is highly device dependent since IoT devices are in a resource-constrained environment. Conﬁdentiality should be used to preserve the exchanged data in the whole architecture of the middleware. It can also ensure that the data inside an entity is protected from unauthorized access.
* ***Integrity*:** This requirement ensures that an exchanged message has not been changed during the transmission by an unauthorized part through data validation and veriﬁcation. IoT entities exchange critical data with other entities, which put forward stringent demand that sensed, stored and transmitted data must not be tampered either maliciously or accidentally. Integrity protection of device data is crucial for designing reliable and dependable IoT applications. This is ensured with message authentication codes (MAC) using one-way hash functions. The selection of MAC technique again depends on device capabilities. Integrity can also be used to protect data stored in entities.
* ***Availability*:** It is important that IoT services be available from anywhere at any time in order to provide information continuously. There is no single security protocol that can satisfy this property. However, different pragmatic measures can be taken to ensure the availability.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Potential Attacks | AUTa | AACb | CCPc | CONd | INTe | AVAf |
| Man-in-the-Middle |  |  | **✓X** | **X** | **X** |  |
| Message Modiﬁcation |  |  | **✓X** |  | **X** |  |
| Masquerading | **X✓** | **X** |  |  | **X✓** |  |
| Authentication Attacks | **X✓** | **X** |  |  |  |  |
| DoS Attacks |  |  |  |  |  | **X✓** |
| Physical Attack | **X** |  |  |  | **X** |  |
| Message Replay |  |  | **X** |  | **X** |  |
| Eavesdropping |  |  | **X✓** | **X** |  |  |
| Traﬃc Analysis |  |  | **X** | **X** |  |  |

Table ‑ : Relationship between attacks and requirements. SOURCE: ResearchGate White Paper, Figure 3, page 409

### API Security Risks in 5G for IoT

Risks to REST APIs are nothing new. There are numerous examples of past exploitation of failure to execute good software programming practices. The following sections describe examples of specific API security risks that should be considered in the development of the REST APIs that are developed to support IoT service development.

#### REST API Security Risk #1: HTTPS Protected API Without Any Authentication

Providing an API using HTTPS is familiar to most developers already. But using an API not having any authentication for personalized services can be tricky as the Nissan Leaf Example tells us. Their API used a Vehicle Number as identifier to allow actions like turning on the AC or reading the cars logbook. Unfortunately, those vehicle numbers were visible on the front window of all those cars (so if your neighbor had one, it was easy to get it) and were in the same number range. This issue resulted in Nissan shutting down this interesting service due to their inability to patch the service quickly.

#### REST API Security Risk #2: No Rate Limiting or Throttling Implemented

By failure of an Android App, the National Weather Service had to shut down the service for some time. The API was not throttled nor limited so the traffic peak directly hit the backend. A good practice is to enforce a system wide quota so that the backend cannot be overloaded. An even better practice is to set up an app and/or user-based quota. This can be easily rolled out with the help of a modern API Developer Portal. It’s also important to monitor overload situations closely and having implemented API keys or OAuth to help identify the root cause more easily.

#### REST API Security Risk #3: Unprotected Identity and Keys

An attack to Buffer occurred in 2013 where tokens and keys were stolen resulting in hackers being able to send spam into social media networks using accounts using Buffer. **OAuth / OpenID Connect** are used to grant access to services on behalf of the user. These tokens give access for a longer period of time without handing over the password of the user to the app. This is a good thing but users are usually not aware of this and struggle to understand the difference between password change and revoking access. The full disclosure Buffer actually helped understanding the topic in more detail. In a few words, if you store any kind of identity or identity token somewhere, make sure that the system and APIs are protected and monitored very carefully. This starts not only with the API itself but also where the data is stored, backed up, etc.

#### REST API Security Risk #4: Unencrypted Payload

There are still some apps – even if seldom – using unencrypted connections or payloads. Listening to traffic and API calls of websites and mobile apps is actually easier than you think. In a browser, it’s just an F12 keystroke to show the browser developer view. Tools like Charles Proxy or Fiddler make it even easier to listen into the connection between the mobile App and the API. A correct use of HTTPS (see below) and if needed payload encryption can help mitigate the risk of exposure.

#### REST API Security Risk #5: Incorrect Use of HTTPS

As explained in the previous paragraph, using HTTPS does not prevent the connection from being intercepted. To prevent this from happening, Certificate Key Pinning needs to be implemented. Especially on mobile apps, there is native mobile Operating System (OS) support for protecting certificates and use of Mutual Secure Sockets Layer (SSL). All sensitive data like end-user credentials, API Keys, etc. should be sent over this Mutual SSL connection only.

#### REST API Security Risk #6: Weak API Keys

API keys are a good way to identify the consuming app of an API. Unfortunately, sometimes the key is sent as part of the URL which makes it appear in proxy logs or other places and easy to copy. Even if HTTPS is used correctly (see above), it can be sniffed. A good way to go around this is the use of Amazon Web Services (AWS) Style API Keys. Putting API Keys in the message header (which is usually not logged on public hotspots) is a good practice and can even go further by using two keys:

1. Secret Key ID to perform Hash-based Message Authentication Code (HMAC) signing (with detection of replay attacks)
2. Access Key ID to identify the client

#### REST API Security Risk #7: Logic and Security in the Wrong Place

Last but not least, security and implementation need to be in the right place. Looking at this example from the Dominos application which allowed faking a payment shows that it’s a good practice to put critical logic like payments or approvals into a secure place, usually not in the client app.[[19]](#footnote-19)

There are tools and procedures available to solve these issues. API Gateways and API Management systems provide protection against:

* Unauthorized Access
* Parameter manipulation and Data harvesting
* Network eaves dropping
* Disclosure of sensitive customer data
* Message replay
* Virus insertion

Besides platforms, there are best practices in API development and enforcement that can mitigate many of these threats. These are explored further in the sections below.

### API Security Best Practices to Mitigate Risks to IoT in 5G[[20]](#footnote-20) [[21]](#footnote-21)

REST (or REpresentational State Transfer) is a means of expressing specific entities in a system by URL path elements. REST is not architecture, it is an architectural style to build services on top of the Web. REST allows interaction with a web-based system via simplified URLs rather than complex request body or POST parameters to request specific items from the system. REST-style service design is recommended whenever possible in 5G. Much research has been done on common best practices for securing RESTful APIs.

#### Authorization

* ***Protect HTTP methods:*** RESTful API often use GET (read), POST (create), PUT (replace/update) and DELETE (to delete a record). Not all of these are valid choices for every single resource collection, user, or action. Make sure the incoming HTTP method is valid for the session token/API key and associated resource collection, action, and record. For example, if you have a RESTful API for a library, it's not okay to allow anonymous users to DELETE book catalog entries, but it's fine for them to GET a book catalog entry. On the other hand, for the librarian, both of these are valid uses.
* ***Whitelist allowable methods:*** It is common with RESTful services to allow multiple methods for a given URL for different operations on that entity. For example, a GET request might read the entity while PUT would update an existing entity, POST would create a new entity, and DELETE would delete an existing entity. It is important for the service to properly restrict the allowable verbs such that only the allowed verbs would work, while all others would return a proper response code (for example, a 403 Forbidden).
* ***Protect privileged actions and sensitive resource collections:*** Not every user has a right to every web service. This is vital to protect administrative web services from misuse. The session token or API key should be sent along as a cookie or body parameter to ensure that privileged collections or actions are properly protected from unauthorized use.
* ***Protect against cross-site request forgery:*** For resources exposed by RESTful web services, it's important to make sure any PUT, POST, and DELETE request is protected from Cross-Site Request Forgery (CSRF). Typically, one would use a token-based approach. CSRF is easily achieved even using random tokens if any Cross-Site Scripting (XSS) exists within your application, so care must be taken to prevent XSS.

#### Input Validation

* ***Assist the user > Reject input > Sanitize (Filtering) > No input validation:*** Help the user input high quality data into web services, such as ensuring a Zip code makes sense for the supplied address, or the date makes sense. If not, reject that input. If they continue on, or it's a text field or some other difficult to validate field, input sanitization is a losing proposition but still better than XSS or Structure Query Language (SQL) injection. If you are already reduced to sanitization or no input validation, make sure output encoding is very strong for the application. Log input validation failures, particularly if the client-side code is going to call web services. The reality is that anyone can call web services, so assume that someone who is performing hundreds of failed input validations per second is up to no good. Consider rate limiting the API to a certain number of requests per hour or day to prevent abuse.
* ***URL Validations:*** Web applications/web services use input from HTTP requests (and occasionally files) to determine how to respond. Attackers can tamper with any part of an HTTP request, including the URL, query string, headers, cookies, form fields, and hidden fields, to try to bypass the site’s security mechanisms. Common names for common input tampering attacks include: forced browsing, command insertion, cross site scripting, buffer overflows, format string attacks, SQL injection, cookie poisoning, and hidden field manipulation.
* ***Validate incoming content-types:***When POSTing or PUTting new data, the client will specify the Content-Type (e.g. application/xml or application/json) of the incoming data. The server should never assume the Content-Type it should always check that the Content-Type header and the content are the same type. A lack of Content-Type header or an unexpected Content-Type header should result in the server rejecting the content with a ***406 Not Acceptable*** response.
* ***Validate response types:*** It is common for REST services to allow multiple response types (e.g. application/xml or application/json), and the client specifies the preferred order of response types by the Accept header in the request. Do NOT simply copy the Accept header to the Content-type header of the response. Reject the request (ideally with a **406 Not Acceptable** response) if the Accept header does not specifically contain one of the allowable types. Because there are many Multipurpose Internet Mail Extensions (MIME) types for the typical response types, it's important to document for clients specifically which MIME types should be used.
* ***Extensible Markup Language (XML) input validation:***XML-based services must ensure that they are protected against common XML based attacks by using secure XML-parsing. This typically means protecting against XML External Entity attacks, XML-signature wrapping etc.

#### Output Encoding

* ***Security Headers:***To make sure the content of a given resources is interpreted correctly by the browser, the server should always send the Content-Type header with the correct Content-Type, and preferably the Content-Type header should include a charset to denote character encoding. The server should also send an “X-Content-Type-Options: nosniff” to make sure the browser does not try to detect a different Content-Type than what is actually sent (can lead to XSS). Additionally, the client should send an “X-Frame-Options: deny” to protect against drag and drop clickjacking attacks in older browsers.
* ***JSON Encoding:*** A key concern with JSON encoders is preventing arbitrary JavaScript remote code execution within the browser or on the server (if using node.js). It is vital to use a proper JSON serializer to encode user-supplied data properly to prevent the execution of user-supplied input on the browser. When inserting values into the browser Document Object Model (DOM), strongly consider using .value/.innerText/.textContent rather than .innerHTML updates, as this protects against simple DOM XSS attacks.
* ***XML Encoding:*** XML should never be built by string concatenation. It should always be constructed using an XML serializer to convert the object into a form that can be readily transported – in this case a proper XML stream. This ensures that the XML content sent to the browser can be parsed and does not contain XML injection.

#### Cryptography

* ***Data in transit:*** Unless the public information is completely read-only, Transport Layer Security (TLS) should be used, particularly where credentials, updates, deletions, and any value transactions are performed. The overhead of TLS is negligible on modern hardware, with a minor latency increase that is more than compensated by safety for the end user. Consider the use of mutually authenticated client-side certificates to provide additional protection for highly privileged web services.
* ***Message Integrity:*** In addition to HTTPS/TLS, JSON Web Token (JWT) is an open standard (RFC 7519) that defines a compact and self-contained way for securely transmitting information between parties as a JSON object. JWT not only ensures the message integrity but also provides authentication of both message sender and receiver. The JWT includes the digital signature hash value of the message body to ensure message integrity during the transmission.

#### HTTP Status Codes

* ***HTTP defines status code:*** When designing a REST API, don't just use 200 for success or 404 for error. Below are general guidelines to consider for each REST API status return code. Proper error handle may help to validate the incoming requests and better identify the potential security risks.
  + - 200 OK - Response to a successful REST API action. The HTTP method can be GET, POST, PUT, PATCH or DELETE.
    - 400 Bad Request - The request is malformed, such as message body format error.
    - 401 Unauthorized - Wrong or no authentication ID/password provided.
    - 403 Forbidden - Used when the authentication succeeded but authenticated user does not have permission to the requested resource.
    - 404 Not Found - When a non-existent resource is requested.
    - 405 Method Not Allowed - The error checking for unexpected HTTP method. For example, the REST API is expecting HTTP GET, but HTTP PUT is used.
    - 429 Too Many Requests - The error is used when there may be DoS attack detected or the request is rejected due to rate limiting.
* 401 vs 403:
  + - 401 “Unauthorized” really means Unauthenticated, “You need valid credentials for me to respond to this request”.
    - 403 “Forbidden” really means Unauthorized, “I understood your credentials, but so sorry, you’re not allowed!”

## Improvements to security in 5G for IoT

5G systems are going to be service-oriented. This implies that there will be a special emphasis on security requirements that stem from the specific features of the differentiated services. For instance, remote health care requires resilient security while IoT requires lightweight security. It is quite reasonable to offer differentiated security to different services.

ATIS is currently working on IoT classifications, designed to aid network providers in determining how best to manage IoT devices. Since not all devices need the same level of service, it is important to understand the type of device and give it a classification.

If differentiated security is offered, then flexible security architecture is needed. Leveraging the benefits of network slicing architecture in 5G can support the separation of specific end-to-end protection for different IoT services. In this approach, the network manages different security capabilities including strength of security algorithms, ways to derive and negotiate secret keys, and mechanisms for protecting confidentiality and integrity. Within a virtual network slice, security capabilities could further be distributed.

IoT requires an evolution of the service framework provided by network operators in order to support the flexibility and diversity of services offered. The risks and mitigation techniques for protecting the network elements and connectivity are discussed elsewhere in this report. The focus of this section has been to address the need to expose the network core to the service provider ecosystem in such a manner that the interfaces are protected.

Specific techniques to protect the SOA from new threats have their origin in the information technology and software communities. These techniques have been developed and improved over the past few years to enable protection for a myriad of applications. The same is true for the best practices identified to protect the RESTful APIs that are recommended for use in 5G networks. The application of these proven techniques will be necessary to harden the middleware and APIs that will be deployed in 5G networks. For example, medical devices would fall under the category “smart health and wellness.” Under this classification there are sub-classes, such as “remote robotic surgery.” These classifications make it easier to assign profiles with policies that can then be applied in their respective service.

# NFV and SDN

SDN defines a network architectural approach rather than a specific product and is based on the physical disaggregation of the network control and data forwarding functions. Separation of these two functions enables the abstraction of applications from the underlying network and enables programmatic control of the network. This decoupling creates a network architecture that is dynamic, manageable, adaptable, and more cost-effective than traditional architectures, making it well-suited for the high-bandwidth, dynamic nature of emerging applications.

The need for SDN was driven by the high-bandwidth, low latency requirements of current applications. Traditional network architecture, defined by a hierarchical tier of switches and routers, is ill-suited to meeting the demands of today’s applications. Additionally, the capital investment required to continually expand and maintain a network comprised of purpose-built hardware and its associated software and support costs are unsustainable.

SDN architectures generally are comprised of three components grouped by functionality:

* SDN Applications are the programs that communicate behaviors and the resources needed within the SDN Controller. Applications include networking management, analytics, or business applications. Additionally, applications could build an abstract view of the network by collecting information from the controller and presenting it for decision-making purposes.
* The SDN Controller is a centralized function that receives instructions or requirements from the SDN Applications, and relays those to the SDN networking components. The controller may also extract information about the network from network devices and present to the applications an abstracted view of the network. The controller is responsible for steering traffic flows to specific network functions and to other networks. It facilitates and manages connectivity in the network. This means it must configure the network infrastructure to accommodate the traffic flows as well.
* SDN Networking Devices execute the data forwarding and processing capabilities according to the routing instructions provided by the SDN Controller.

The other network architecture that will see expanded application in 5G networks is the virtualization of network functions. NFV eliminates the need for custom hardware appliances dedicated to each network function. NFV focuses on three main components of virtualization: compute, storage, and networking. A virtualized network function (VNF) may consist of one or more virtual machines running different software and processes on high volume servers, switches or storage devices, or even within cloud computing infrastructure. This separation of the network function from the hardware device on which it runs is the key to reducing total cost of ownership because VNF’s can now run on lower cost, commercial off the shelf (COTS) hardware, eliminating the need for more expensive, purpose-built hardware.

While this was the main objective and driver for NFV, the reality has become just the opposite. Dedicated hardware is engineered for the specific application it supports, and therefore is optimized for that application. When using COTS for the same application, the performance is far below the engineered system, resulting in a loss in performance as much as 30%. This means that currently more hardware is needed in an NFV environment than would be needed with engineered systems.

However, NFV does deliver on its promise to provide a flexible and elastic network capable of meeting the demands of network traffic in real-time. This has become the new driver for NFV implementations.

The NFV framework consists of three main elements:

* VNFs are the software implementations of network functions that are deployed on a network function virtualized infrastructure (NFVI).
* NFVI is comprised of all the hardware and software components in the environment where VNFs are deployed. The NFVI could span multiple physical locations, and the network connectivity among those locations is generally considered to be a part of the NFV infrastructure.
* Orchestration, or NFV management, is the collection of all the functional blocks, data repositories, reference points and interfaces through which the functional blocks exchange information for managing and orchestrating the NFVI and VNFs. Orchestration provides the capability to instantiate VNF instances, monitor and repair them, and bill for services rendered.

While distinct architectures[[22]](#footnote-22) that can be deployed separately, NFV and SDN can be combined to allow operators to introduce new network functions and add network capacity as needed, rather than waiting for hardware and software releases on legacy platforms. This combination also allows operators to use common hardware in their data centers to support new services.

It also supports the ability to fully implement a new network function based on real-time demands (elasticity) and delete the instance of a network function when it is no longer needed. The hardware can then be repurposed for another instantiation of a network function as needed.

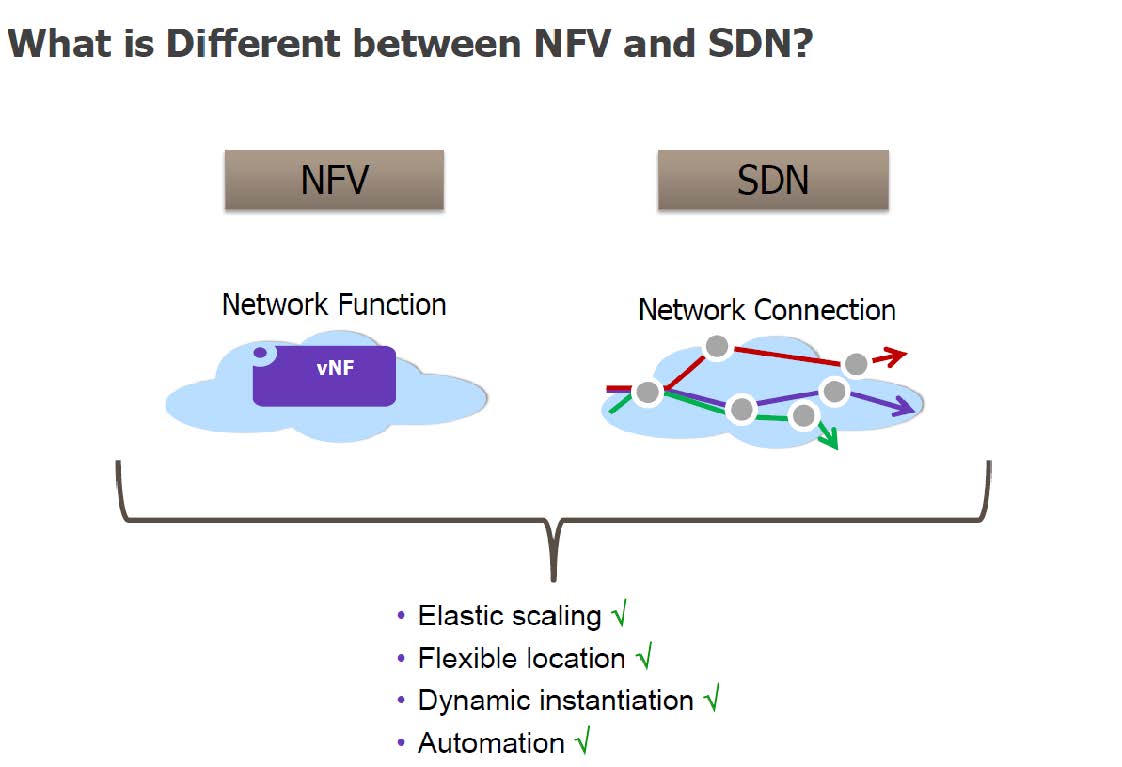


Figure ‑: NFV vs. SDN. SOURCE: FCC TAC Cybersecurity Working Group, White Paper: Considerations for Securing SDN/NFV

The SDN is managed by an SDN controller. The controller is responsible for steering traffic flows to specific network functions and to other networks. It facilitates and manages connectivity in the network. This means it must configure the network infrastructure to accommodate the traffic flows as well.

The 5G architecture is a native NFV/SDN architecture covering aspects ranging from devices, mobile and fixed infrastructure, network functions, value-add services, and the management functions that orchestrate a 5G system.

To realize such a 5G system architecture, the control (C) and user (U) planes are separated, with open interfaces defined between them. By making this separation, the user plane can expand and contract to meet ever-changing demands without affecting the rest of the network functions under the command of the control plane. This also enables the concept of network slicing in accordance with SDN principles. NFV is what allows the network functions to be virtualized (virtual network functions, or VNFs), and orchestrated.

It is important to understand that a VNF does not necessarily have to be instantiated on its own dedicated hardware. VNFs can co-exist on the same server, provided there is isolation between the two VNFs. The hypervisor resident on the hardware is responsible for ensuring this isolation.

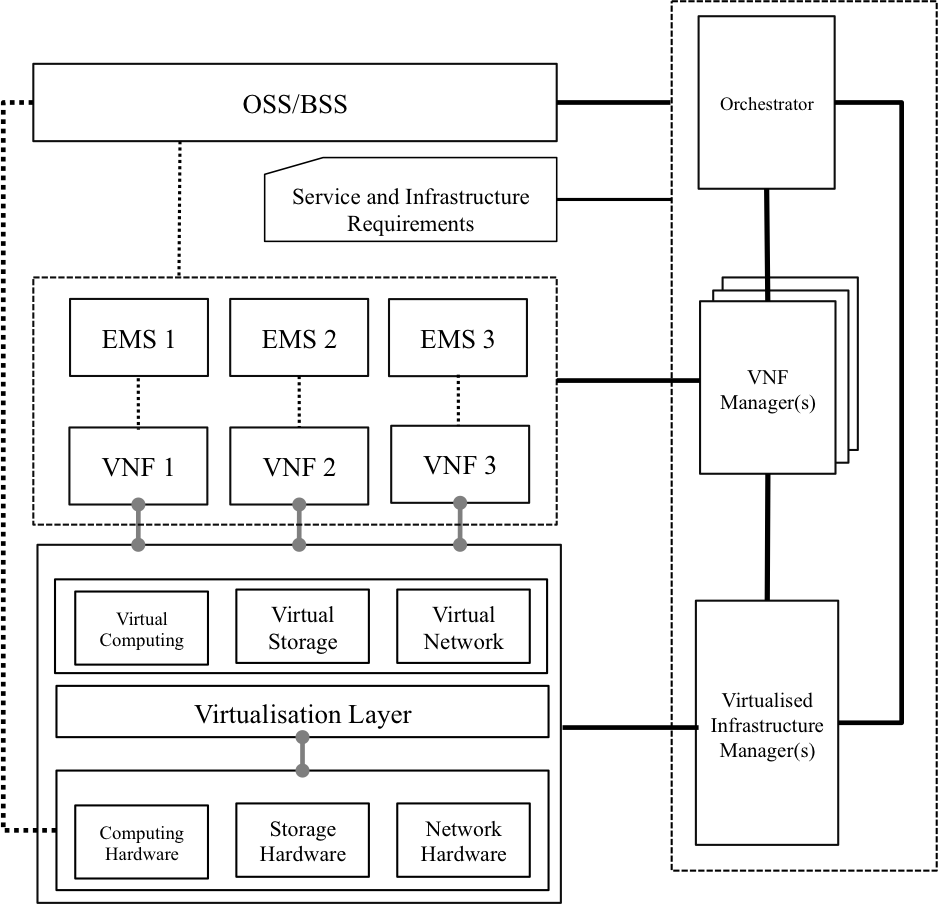


Figure ‑: NFV Architecture. SOURCE: Source: ETSI GS NFV-SWA 001 V1.1.1

While it is possible to use physical isolation between VNFs, this may not always be economical or efficient. One of the principle drivers behind NFV and SDNs has always been the ability to share hardware used across multiple functions, eliminating the need for dedicated hardware and purpose-built hardware.

As seen in the figure above, the OSS/BSS layer interfaces directly with the orchestrator. When it is determined that a new VNF is to be instantiated (based on analytics monitoring network conditions for example) the orchestrator chooses the function to be instantiated from a catalog (provided by the OSS). The OSS provides the template used by the element manager to build and provision the new VNF, based on predefined definitions.

For example, if it is determined that a new instantiation of an HSS is needed because of an increase in network traffic, the OSS provides a template for how the HSS is to be provisioned. This allows for automatic instantiation of new network elements without manual intervention.

The VNF manager is also responsible for ensuring the proper interfaces between the instantiated VNFs are properly defined and provisioned. As a new HSS instance is created, the correct interfaces for the HSS must be defined and implemented between the new HSS instance and the other network functions.

The software used to create a new VNF instance is provided, including any open source or third- party components, as a VNF package. This same software is re-used across multiple VNFs with the same function. It is the responsibility of the vendor providing the software to ensure that all of the components comprising the VNF have been tested and will operate properly, however, it will be the responsibility of the operator to test the VNF on the hardware (or virtual machine) being used in the network. This is a consideration that operators must take into consideration when designing their virtual networks.

Not all network functions can be virtualized. Some may have specific hardware requirements. In these cases, some components of the network function (such as operations, administration, and maintenance [OAM]) may be virtualized, while the rest of the network function is run on dedicated, purpose-built hardware.

When implementing a new VNF on hardware, there are a couple of approaches that can be used. Most will run the network function on a virtual machine, using a hypervisor. The hypervisor emulates certain functions of the hardware, allowing any operating system to run on top of another operating system. Each of the operating systems use their own kernel.

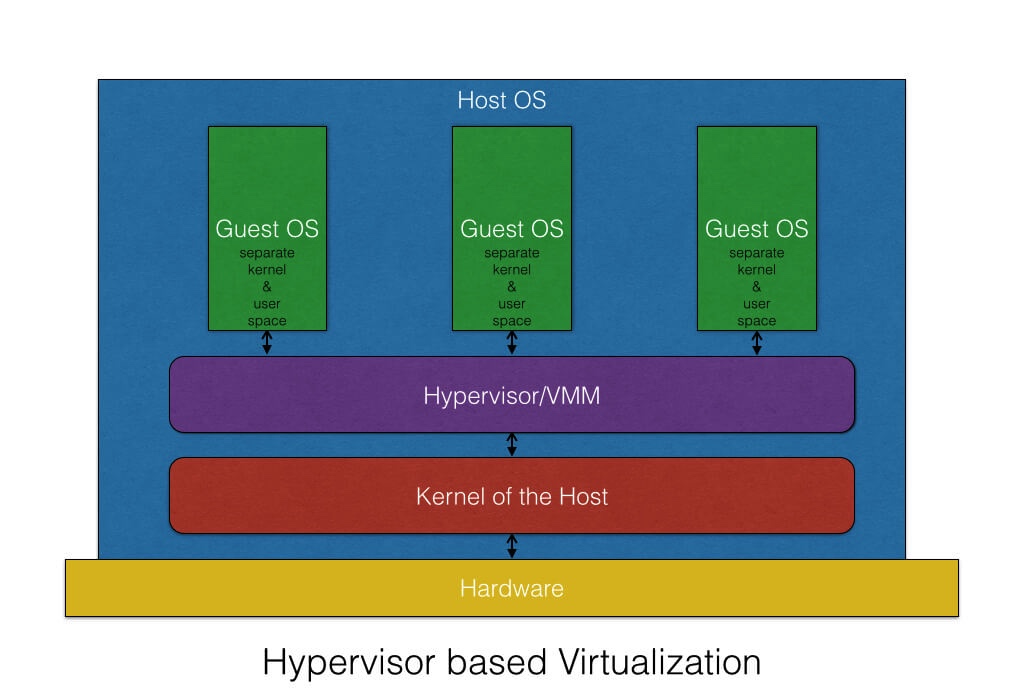


Figure ‑: Hypervisor based virtualization. SOURCE: <https://blog.risingstack.com/operating-system-containers-vs-application-containers/>

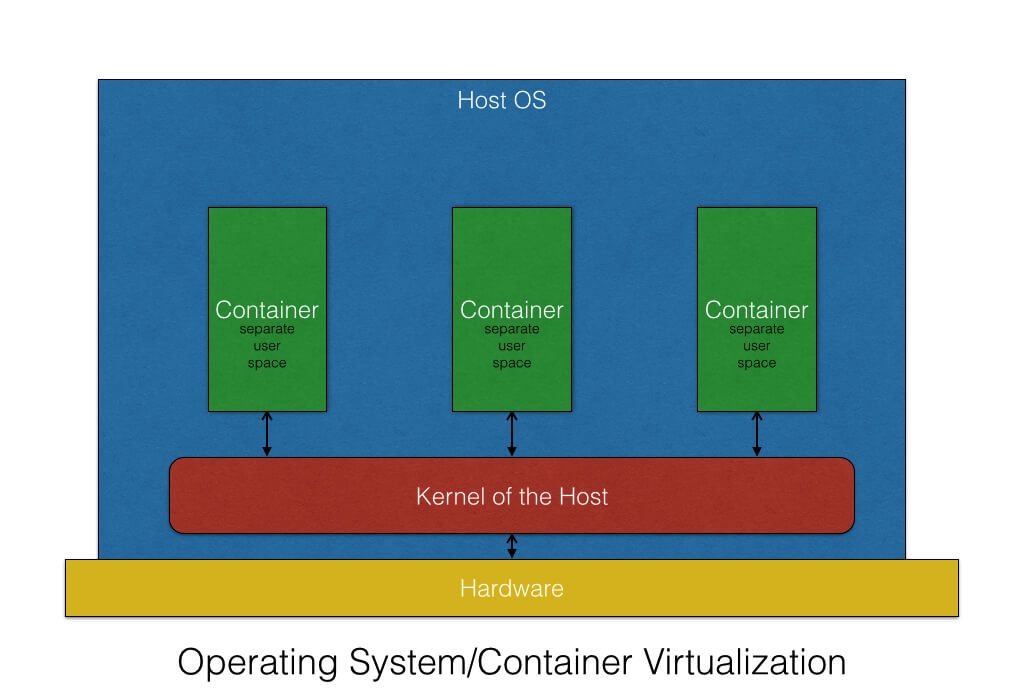


Figure ‑: Operating system/container virtualization. SOURCE: <https://blog.risingstack.com/operating-system-containers-vs-application-containers/>

Containers are different. A container provides a lightweight environment of commonly used resources such as memory, CPU, or storage from the host hardware. All processes inside the container are isolated and are unable to see the processes in any other container. The container shares the same kernel as all other containers, which means it will run faster than a virtual machine using hypervisors.

However, you cannot run a Windows operating system on a Linux host like you can when using a hypervisor. And because the containers all share the same kernel, any compromise of the common kernel exposes all of the containers.[[23]](#footnote-23)

A network function can be dependent on a specific hypervisor, or it can be hypervisor agnostic. Likewise, it could require the use of a container. This will depend on the vendor and the application being provided.[[24]](#footnote-24)

The Management and Orchestration (MANO) layer is an important consideration in mobile networks that becomes even more critical when using NFV and SDN. The specific implementation of MANO standards in the form of cloud manager products is used to automatically deploy and orchestrate VNFs. The industry is cognizant of the risks in NFV/SDN, and their concerns are being addressed in multiple standards bodies.[[25]](#footnote-25)

The industry is cognizant of the risks in NFV/SDN, and their concerns are being addressed in multiple standards bodies.

## Risks of NFV and SDN in 5G

The capability to virtualize network functions and define networks within software brings significant advantages to network operators. 5G networks will be built on a virtualized platform taking advantage of NFV, SDN and Containerization along with Open Network Automation Platform (ONAP). Therefore, the security advantages associated with virtualization, and ONAP will apply to 5G. For example, closed-loop automation based on ONAP along with virtualization’s inherent elasticity feature will be leveraged as a significant 5G security advantage. Those advantages, however, carry with them risks previously unseen in traditional hardware-defined networks.

Because NFV and SDN are still comparatively new architectures, many network operators may still be unfamiliar with the risks inherent in networks more defined by software than hardware. This unfamiliarity in itself introduces new risk.

To take full advantage of the benefits of NFV and SDN, as well as to provide the greatest security for these networks, operators will need new tools to ensure they have full visibility and control of network topology. While these principles have been in place in data centers for some time, they are often new to telecom engineers and planners.

The ability to deploy NFV and SDN on COTS hardware, thereby avoiding the cost of purpose-built appliances such as routers and firewalls, represents one of the major attractions of network deployments using NFV and SDN.

Unlike traditional networks in which security models are relatively static, in the future, orchestration of NFV / SDN network topology could introduce continually changing workloads. Reacting to such a dynamic environment requires that NFV and SDN rely on centralized orchestration to manage workloads, creating new virtual machines as network demand dictates. The hypervisor, that function whose role it is to ensure isolation of virtualized functions, introduces attack vectors in a virtualized network, which could be mitigated by NFV/SDN dynamic security controls. Compromising the hypervisor exposes the network to orchestration exploits, SDN controller exploits, data exfiltration or destruction, and malicious configuration attempts.

Elastic network boundaries are a totally new concept introduced by NFV and SDN. Traditional network models are fixed, and network boundaries are easily identified by the existence of hardware at the network edge. NFV and SDN network boundaries, on the other hand, can change based on the dynamic nature of workloads. This means that network topologies might change, and the automatic configuration capability of network orchestrators has introduced new vulnerabilities to an NFV/SDN network. Compromising the network orchestrator might allow parties to implement malicious network configurations, altering network parameters and moving a VNF to an unauthorized location. This could lead to a regulatory compliance failure in some cases. Additionally, this elasticity introduces the opportunity for amplification attacks when new instances of network functions can be orchestrated based on network dynamics or attacks on orchestrators or controllers. If an orchestrator were compromised, new virtual machines (VMs) could be instantiated, with the potential for a resultant flooding of network resources. This is another area where the dynamic security controls introduced by NFV/SDN will improve on security.

In addition to these new attack vectors, NFV/SDN introduces the increased use of open source software. Open source software is often not subjected to the rigorous patching and update disciplines to which traditional network element software – that software with which many network operators are most familiar – is subject. This places a heavier burden for security on the operator and moves away from regularly scheduled updates from vendors to a model where software updates are continuous, based on need. While a security advantage, this can also be a risk. Open source software benefits and risks are described in the Open Source section of this report.

SDN represents a departure from networks with known topologies that were designed over time, and implemented in a methodical process. SDN networks are dynamic, and ever-changing based on the real-time needs of the network. This means network operations will need new tools to ensure they have complete visibility and control of the ever-changing network topology. While these principles have been in place in data centers for some time, they are new to telecom engineers and planners. Traditional models based on purpose-built hardware and software maintained with rigorous patching and management disciplines will be replaced by highly flexible, SDNs that are agile enough to meet the demands of future networks and subject to continuous update. Taking full advantage of these capabilities will place a greater burden on both the operator and the vendor to manage the risks associated with that greater agility. The figure below shows the various methods that are needed for securing the SDN.

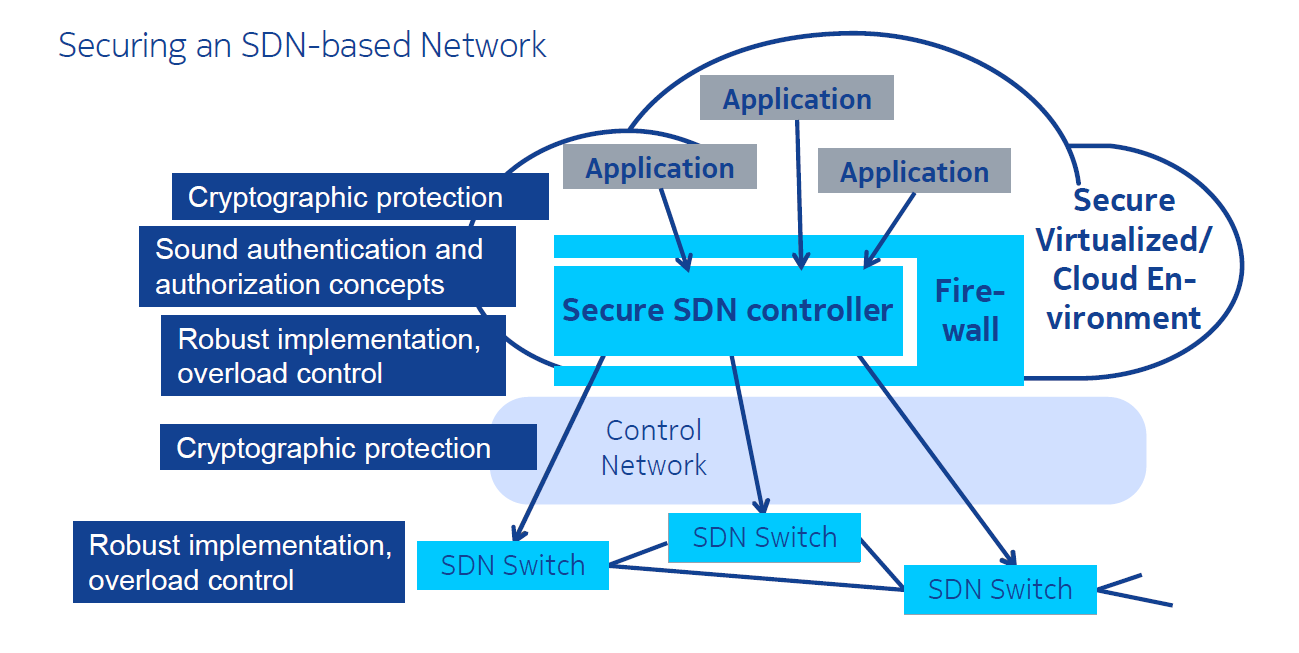


Figure ‑: Securing the SDN. SOURCE: Peter Schneider, Nokia.

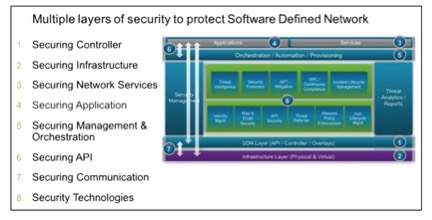


Figure ‑: Layers of security in SDN. SOURCE: Mike Geller, Cisco.

## Threat assessment of NFV and SDN in 5G

NFV and SDN present new attack vectors that will be unfamiliar to operators who have not implemented this technology before. These vectors, however, are not new to 5G but are a result of the virtualization of network functions and the deployment of automated controls for services orchestration that are necessary in these emerging architectures. The benefit is that operators currently deploying this architecture in the 4G LTE environment are already developing experience that will benefit their deployment and operation of the 5G network.

An attack on a hypervisor could have a major impact on the network, depending on the deployment of the hypervisor.

For example, if the hypervisor is supporting multiple critical network functions on the same hardware, a compromised hypervisor could result in a DoS on all of the network functions supported by the hypervisor. This is where isolation plays an important role, but physical isolation is not always the best choice, because it eliminates the advantages of NFV/SDN. Physical and logical isolation of network functions must be carefully weighed to ensure maximum security of the network function, while still realizing the benefits of NFV/SDN.

A lack of skillsets in virtualized networks represents perhaps the largest threat in NFV/SDN networks. This is new technology for telecommunication networks, and even though NFV/SDN has been around the IT world for some time, the standards for telecommunications NFV/SDN are new.

This means even a seasoned IT professional with experience in NFV/SDN will have to learn the new standards for NFV/SDN and its use in a telecommunications network. There are many important differences between the two types of networks and understanding these differences will be critical to securely maintaining NFV/SDN networks.

Most operators will have legacy monitoring systems in their 3G/4G networks. These monitoring systems are designed to collect network data from various standardized interfaces in the network using probes located in strategic points, such as at major routers.

With NFV/SDN, these probes will no longer be effective. They will not can collect data between virtual machines and given the orchestration of new instantiations of network functions, they will not have visibility to newly instantiated network functions. Consideration needs to be given to how visibility will be enabled in a virtual network, and between VNFs. This can be accomplished through the virtual switch itself, or software agents that are instantiated with the VNF.

Since there lacks any industry standard definition for how virtual monitoring in telecom networks should be implemented, operators will be dependent on their vendors to define how virtual monitoring systems will work in their specific NFV/SDN networks. The lack of standards will result in proprietary systems that will lack interoperability. Operators will need to work with their vendors to ensure an approach that can later be aligned with industry standards for monitoring in a virtual network.

Security requirements in SDN / NFV network architecture differ significantly from the manner in which conventional networks have been secured. While conventional networks were generally considered to be secure when the edge elements were secure, the architecture of networks that are software defined and centrally controlled, and which are comprised of off-the-shelf hardware from a variety of vendors exposes those networks to new attacks vectors.

General classifications of potential attacks on software-defined networks include the following (although dependent on individual network configuration, all risk vectors may not exist in every software defined or virtualized network):

* **Network Manipulation.** This attack occurs on the control plane, when an attacker can compromise the SDN controller. Due to its basic architecture, the SDN Controller is effectively the “brain” of the network, containing the capability to program traffic flows within the network. Therefore, attacks on the Controller are the most severe threats to the SDN architecture. Once gaining access to the SDN controller, the attacker can produce false network data and initiate attacks on other network elements or the entire network. A compromised controller can give “root-like” access, enabling an attacker to configure virtualized network elements under its control, leading to data loss or further loss of network security.
* **Traffic Diversion.** This is an attack on the network element at the data plane. It allows the attacker to redirect traffic flows and allows eavesdropping.
* **Side Channel Attack.** This attack on the network elements, generally occurring in the data plane, results from the collection of externally observable network behavior by an attacker during normal network operations. Not requiring a connection to the network, the attacker might use information such as – for example – network timing, to determine the length of time it takes the network to establish a connection. With that information, the attacker could determine whether or not a flow rule exists.
* **Application Manipulation.** Because SDN architecture is by nature defined in software, the network can no longer be defended solely by physical topology, and network firewalls generally are not designed to stop attacks on applications. Further, software applications are becoming increasingly complex, and security is not always properly built into applications. For these reasons, a successful attack on an application could provide the attacker with the ability to impact other portions of the network.
* **Denial of Service.** Perhaps the most common of attacks, this could impact all parts of an SDN. An attacker could flood the network with malicious messages, potentially cause a reduction or complete disruption of service to the network.
* **Access Resolution Protocol (ARP) Spoofing (“Man-in-the Middle Attack”).** ARP spoofing is a type of attack in which a malicious actor sends falsified ARP messages over a network, resulting in the linking of an attacker’s MAC address with the IP address of a legitimate computer on the network. Once the attacker’s MAC address is connected to an authentic IP address, the attacker will begin receiving any data that is intended for that IP address. ARP spoofing can enable malicious parties to intercept, modify or even stop data in-transit. ARP spoofing attacks can only occur on local area networks that utilize the Address Resolution Protocol and are not specific to NFV and SDN networks.
* **API Exploitation.** Initially attacking the northbound interface, this attack could permit destruction or modification of data flows and could enable the unauthorized disclosure of data. It is important to recognize that API’s may contain vulnerabilities that would allow an attacker access to other network elements.
* **Traffic Sniffing.** An SDN hacker can take advantage of any unencrypted communication interface to intercept or interfere with traffic to and from a central controller or network element.

In summary, NFV security should provide sound, robust implementations of the virtualization layer and the overall cloud platform

* Sound robust security aware implementation of the VNFs
* Integrity assurance for both platform and VNFs

In virtual networks, isolation needs to be considered for critical network functions. Isolation can be implemented using logical or physical isolation. Even traffic isolation can be important for certain types of network traffic.

Also, virtual firewalls can provide perimeter security and network traffic filtering in addition to logically or physically separated security zones. Encryption should be used when moving traffic from one system to long-term storage

For securing an SDN-based network a secure virtualized environment should include the protections for the SDN controller like cryptography, sound authentication and authorization and robust implementation & overload protection. The control network should use cryptographic protection and the SDN switch should be protected using robust implementation and overload control that implies high-availability implementations with redundancy and fail-over capabilities.

## Improvements in 5G for Securing NFV and SDN

While all potential security threats of NFV and SDN pose some risk and must carefully be mitigated, on the positive side, SDN and the cloud may also bring security advantages for 5G networks. As an example, flexible, scalable allocation of resources to a VNF may help significantly in overcoming certain DoS attacks, which are otherwise difficult to counter. The programmability provided by SDN may allow deploying security solutions flexibly and efficiently, running as applications on top of SDN controllers. Finally, both NFV and SDN may support the automation of tasks in 5G systems that otherwise need to be done in a more manual style and thus could be more prone to human errors inducing security vulnerabilities.

# Open-Source 5G Software Platforms

Open source software includes operating systems, applications, and programs in which the source code is published and made available to the public, enabling anyone to copy, modify and redistribute that code without paying royalties or fees. Open source “products” typically evolve through community cooperation among individual programmers as well as large companies. An open source license permits anybody in the community to study, change and distribute the software for free and for any purpose.[[26]](#footnote-26)

|  |  |
| --- | --- |
| Healthcare Health Tech Life Sciences | 46% |
| Retail & E-Commerce | 41% |
| Cybersecurity | 41% |
| Big Data AI BI Machine Learning | 38% |
| Enterprise Software/SaaS | 37% |
| Internet and Mobile Apps | 37% |
| Other (EdTech MarketingTech IoT) | 36% |
| Internet & Software Infrastructure | 35% |
| Manufacturing Industrials Robotics | 35% |
| Computer Hardware & Semiconductors | 34% |
| Telecommunications & Wireless | 31% |
| Virtual Reality Gaming Entertainment Media | 28% |
| Financial Services & FinTech | 28% |
| Energy & CleanTech | 27% |

Table ‑: Percentage of open source by industry SOURCE: Black Duck

Since open source software first entered the market, software vendors have incorporated this software into their platforms at every level. Open source code is widely used in telecommunications systems everywhere. Operating systems, code libraries, middleware, and much more are often open source. Vendors use open source as a foundation on which they can develop their own software, without having to expend resources for commonly used platforms. According to a Black Duck report, open source software was found in 96% of commercial applications that they scanned. The open source code made up 36% of the applications code base.

Open source software is embedded into the majority of solutions deployed in wireless communication platforms. One example of this is Linux. While there are vendor-specific variants of the Linux operating system they all incorporate significant portions of the core open source Linux. Also, due to the open source licensing terms these vendor-specific Linux variants are also open source.

One of the fundamental differences between open source software and proprietary software is that the source code of open source software must be made available with the software. This does not mean that the source code must be physically delivered with the software, just that it must be available at a freely accessible location.

Commercially available software, or proprietary software, doesn’t give access to its source code because the software is considered the intellectual property of the developer. As a result, users often pay for licensing the intellectual property or software. In comparison, open source software is considered shared intellectual property among all contributors that have helped develop or alter it.

One of the key advantages of using open source is that there are many more developers reviewing and testing the code, and the more popular and widely used the open source software, the more developers reviewing the code and finding flaws or and bugs.

This also means the open source software quality gets improved as more developers contribute to the open source projects. When a bug is found, the code can be quickly patched without having to wait for a vendor to issue a patch in a maintenance release. However, vendors of wireless communication platforms still need to incorporate, test and validate the open source changes and release them to their customers.

Potential advantages of open source software:

* Reduces the cost and effort to produce common functionality.
* Larger numbers of developers producing and maintaining more popular functionality.
* Provides for faster distribution of bug fixes.

Potential disadvantages of open source software:

* Usually not subjected to any formal review, validation or verification processes
* May provide a higher opportunity for malicious code injection.
* Open source software code equally available for study by developers and potential attackers.
* Widespread use of open source software creates a larger base of exploit targets (as opposed to potentially smaller bases of vendor proprietary software)

## Risks of Open Source in 5G

Open source software is incorporated into applications in many ways, and often an operator will not know where open source is used. When open source is used as the foundation for a vendors’ product, any vulnerabilities could threaten the integrity of the vendors’ solution. Open source software provides attackers with a target-rich environment because of its wide spread use.

This means vendors must ensure they have mechanisms in place to monitor Common Vulnerabilities and Exposures (CVEs) against any open source software components that they may use in their own products. Vendors must test and perform security assurance assessments on all open source software and bug fixes. Vulnerabilities such as Heartbleed were exploits that targeted open source software vulnerabilities, threatening systems using the open source code.

According to the Black Duck report, 67% of vulnerabilities discovered in open source code were known for more than four years. 52.6% of vulnerabilities were considered as high-severity by the National Institute of Standards and Technology (NIST). Open source vulnerabilities are published on sites such as the National Vulnerability Database (NVD) and are public documents. Network operators should be monitoring this as well and should understand where their vendors are using open source.

Use of open source will continue to increase as vendors rely on open source software to speed delivery of new solutions. Open source software can be viewed as being analogous to corporations outsourcing functions not related to their core competencies. This introduces a new set of security challenges in terms of keeping a consistent and coherent assurance of security-by-design, and prevention of resulting security flaws.

To compound this issue, asking vendors to disclose the open source components used in their products may disclose more vulnerabilities and add to the risk. Note that many times while there may be a vulnerability in a specific software component, that vulnerability may only exist as a stand-alone component, and may be nullified when incorporated into the vendors’ solution (through middleware where the open source component is isolated, for example). Care must be given in how open source components are disclosed to prevent exposure.

## Threat Assessment for Open Source in 5G

While the use of open source offers benefits to enterprises and development teams in terms of time to market, cost and reliability, it also can be the source of vulnerabilities that pose significant risk to application security. Many development teams rely on open source software to accelerate delivery of digital innovation. Both traditional and agile development processes frequently incorporate the use of prebuilt reusable open source software components. As a result, some organizations may not have accurate inventories of open source software dependencies used by their different applications, or a. process to receive and manage notifications concerning discovered vulnerabilities or available patches from the community supporting the open source.

## Threat Mitigation for Open Source in 5G

For larger enterprises with multiple and vast repositories of code, identification of all of the applications where open source vulnerabilities may exist can be difficult. In order to address the identification and mitigation challenge requires an intentional effort that includes activities such as code inspection, dynamic security scanning and vulnerability testing. These are the same techniques that should be applied to all software code repositories, whether open source or not. There are also enterprise specific products that offer a complete end-to-end solution for third party components and supply chain management with features such as licensing, security, inventory, and policy enforcement. These products are offered by vendors such as Black Duck Software, Sonatype Nexus, and Protecode, to name a few.

Most organizations search the CVE and NIST Vulnerability Database for vulnerability information, but these sources provide little information on open-source vulnerabilities. Information on open-source vulnerabilities is distributed among so many different sources that it is hard to track it. To address the risk of open source vulnerabilities in the software supply chain, groups such as PCI and Open Web Application Security Project (OWASP) have specific controls and policy in place to govern the use of open source components. Other security repositories exist including the Node Security Project for JavaScript/Node.js specific vulnerabilities and Rubysec for Ruby specific vulnerabilities. However, there are still many open source projects and ecosystems that are not well covered.

An entire market of open source and commercial tools has emerged over the years to tackle this problem as a result. These tools vary in approach and capabilities, and some are open-source themselves. Most of these tools use the NIST NVD as a starting point for sourcing open source software vulnerabilities. Each tool is then enhanced with usability features and/or additional data sourcing for improved functionality. A sample of these tools follows[[27]](#footnote-27):

### Node Security Project (NSP)

The NSP is known for its work on Node.js modules and npm dependencies. It also provides tools that scan for dependencies and find vulnerabilities using public vulnerability databases such as the NIST NVD as well as its own database, which it builds from the scans it does on npm modules.

### RetireJS

RetireJS is an open-source, JavaScript-specific dependency checker. The project is primarily focused on ease of use. That's why it has multiple components, including a command-line scanner and plugins for Grunt, Gulp, Chrome, Firefox, ZAP, and Burp. RetireJS retrieves its vulnerability information from the NIST NVD as well as a multitude of other sources, including mailing lists, bug-tracking systems, and blogs for popular JavaScript projects. RetireJS also made a site-checking service available to JS developers who want to find out if they're using a JavaScript library with known vulnerabilities.

### OSSIndex

OSSIndex supports several technologies. It extracts dependency information from npm, Nuget, Maven Central Repository, Bower, Chocolatey, and MSI (which means it's covering the JavaScript, .NET/C#, and Java ecosystems). OSSIndex also provides a vulnerability API for free. OSSIndex currently retrieves its vulnerability information from the NIST NVD.

### Dependency-Check[[28]](#footnote-28)

Dependency-check is an open-source command line tool from OWASP that is well maintained. It can be used in a stand-alone mode as well as in build tools. Dependency-check supports Java, .NET, JavaScript, and Ruby. The tool retrieves its vulnerability information strictly from the NIST NVD.

### Bundler-Audit[[29]](#footnote-29)

Bundler-audit is an open-source, command-line dependency checker focused on Ruby Bundler. This project retrieves its vulnerability information from the NIST NVD and RubySec, which is a Ruby vulnerability database.

### Hakiri[[30]](#footnote-30)

Hakiri is a commercial tool that offers dependency checking for Ruby and Rails-based GitHub projects using static code analysis. It offers free plans for public open-source projects and paid plans for private projects. It uses NVD and the Ruby Advisory Database.

### Snyk[[31]](#footnote-31)

Snyk is a commercial service that focuses on JavaScript npm dependencies. Snyk goes beyond the ability to detect known vulnerabilities in JavaScript projects by also helping users fix these issues using guided upgrades and open-source patches that Snyk creates. Snyk has its own vulnerability database, which gets its data from the NIST NVD and the NSP. Snyk's focus is on scaling known vulnerability handling across the entire organization and its teams, with better collaboration tools and tighter GitHub integrations.

### Gemnasium[[32]](#footnote-32)

Gemnasium is a commercial tool with free starting plans. Gemnasium has its own database that draws from several sources. While the vulnerabilities are reviewed manually on a daily basis, advisories are not automatically published. Gemnasium provides a unique auto-update feature that uses a special algorithm to test smart combinations of dependency sets instead of testing all the combinations, which saves time. Gemnasium supports Ruby, npm (JavaScript), PHP, Python, and Bower (JavaScript). Another unique offering from Gemnasium is its Slack integration—users are notified through Slack in real time as soon as an advisory is detected.

### SRC:CLR[[33]](#footnote-33)

Source Clear is a commercial tool with its own database, which leverages the NIST NVD, but it also retrieves vulnerability information from mailing lists and several other sources. It offers a multitude of plugins to several Integrated Drive Electronics (IDEs), deployment systems, and source repositories, as well as a command-line interface. Finally, Source Clear is using "vulnerable methods identification," which is a way to figure out whether a vulnerability found in a dependency is actually being used by the application. It's a feature that dramatically reduces false positives and gives developers detailed target reports for the vulnerabilities that matter.

## Improvements in Open Source as a Result of 5G

As the core network leverages more instances of the traditional Information Technology platforms to enable all services, the use of open source software by mobile network operators will inevitably increase. While this will not drive specific changes in best practices related to the use of open source software, the community will benefit from the expanded use that will result from the addition of companies in the 5G ecosystem. As a result, it is likely that the increased visibility into open source software will result in improvements in vulnerability detection, reporting, and patching.

# Supply Chain in 5G

There has been much debate over the risks introduced through the supply chain. This is especially true in current geo-political climates where there has been additional scrutiny placed on goods from specific nation states.

Regardless of the origin, supply chain risks should be considered for all enterprises. However, as stated by the ISA in a report on supply chain,

“A broad, holistic approach to risk management is required rather than a wholesale condemnation of foreign products and services.”[[34]](#footnote-34)

The working group took this same view as it developed this section of the report. This section will look into the risks introduced through the supply chain and discuss best practices available to enterprises to manage their supply-chain risks. These best practices in use today are directly applicable to the emerging 5G supply chain.

## Supply Chain Background

The 5G supply chain will leverage the existing global supply chain to deliver the next generation of wireless technology performance, capacity and services. Over the past few decades’ forces have reshaped the wireless telecommunications supply chain. As the industry has moved from delivering phone calls to delivering the Internet, major shifts have given birth to a global marketplace, which in turn has resulted in a global supply chain.

Supply chains for wireless telecommunications are complex and the complexity will carry over to 5G. They include development of intellectual property and standards; fabrication of components and chips; assembly and test of devices; development of software and firmware; acquisition, installation, and management of devices and operational networks; and the data and services that operate over those networks. Competing in a global marketplace drives where and how each portion of this supply chain is performed. Supply chain is comprised of every provider that has a relationship with an organization.

5G supply chains are complex and should be viewed on a global scale. An example of wireless supply chain complexity is that of the modern smartphone. It is reported that for the smartphone one can expect that on the order of 700 suppliers from 30 countries may provide components. While the device may be designed in the US, component technologies may come from all over the world and assembled in China – cameras from Japan, displays from Korea, wireless chips from the US, and computer processors from Taiwan.

The best approach for tackling this challenge is thorough Supply Chain Risk Management (SCRM). Through SCRM, areas of risk are identified and prioritized. Specific concerns about particular products or vendors are identified and risk management plans are developed to address concerns.

In this section, we intend to discuss the functional description of the 5G supply chain, general critical infrastructure components and themes important to note, risks that currently result/may result within the global supply chain body, and recommended standards and potential best practices that are seen to mitigate the highlighted risks within supply chain in the 5G environment.

We do not intend to specifically evaluate any suppliers and nation states, identify or address current government policies, and assess specific suppliers within the 5G supply chain environment.

## 5G Supply Chain Ecosystem

Multi stakeholder networks have continued to evolve as supply chains have become increasingly strategic, complex and global. A supply chain ecosystem entails complex international networks of suppliers, partners, regulators and customers – together collaborating to ensure the efficient, effective and competitive movement of products, services, information and funds around the world. With such a multidimensional ecosystem, the supply chain process is external to the communications, hardware, and software stakeholders and “visibility” of the chain becomes obscure to monitor. It requires a broad, inclusive approach allowing organizations to identify their place within the supply chain and map their cyber dependencies and vulnerabilities and will rapidly become more granular in the 5G era.

The established cellular industry ecosystem can be summarized as follows;

* **Component Manufacturers:** At the beginning of the chain, are the component manufacturers that produce billions of components used in the production of hardware and software components used in mobile devices and communication systems. Note that components can be either software or hardware.
* **Vendors:** The above components are assembled using components from in house or third party manufacturers. This vendor’s ecosystem is utilized to build the communication system.
* **Network Operators:** These companies own the spectrum and telecommunication infrastructure and provide the service to consumers. They procure from their vendor supply chain and sell to consumers via stores and/or online.
* **Service Providers:** The companies that provides services that are made available through the communications infrastructure provided by network operators is referred to as service providers.
* **Users:** Finally, there are the end users who consume the services that utilize the networking technologies. These are consumers as well as business/industries.

In the 5G supply chain ecosystem, there will be an expanded role for software developers independent of the vendors associated with the traditional cellular industry supply chain. As a result, hardware and software need to each be considered specifically within the threat and risk mitigation planning of organizations. For example, hardware vendors should be able to identify and validate sourcing of components that are included in a platform. Similarly, software vendors should be able to identify and validate open source software libraries integrated into their solutions. Further, some vendor’s offer outsourced managed operations further complicating the reach of the supply chain. The essentials of good supplier management and SCRM procedures will become even more critical as a result of the expanded in the supply chain.

Industry best practices are described in NIST standards and guidelines as well as other applicable national and international standards and best practices. They include:

* Ensuring that organizations understand the cost and scheduling constraints of implementing supply chain risk management
* Integrating information security requirements into the acquisition process;
* Using applicable baseline security controls as one of the sources for security requirements;
* Ensuring a robust software quality control process;
* Establishing multiple sources, e.g., delivery routes, for critical system elements.

A formal program and process, including dedicated resources, may be used to reach a company defined baseline.

Implementing every foundational practice may be challenging for small- and medium- sized organizations (SMBs) who may, for example, lack the purchasing power to influence supplier behavior, or whose organizational structure does not lend itself to implement a formal program and process to reach a pre-determined and specified level of maturity. The best practices referenced in this section are, however, intended to provide sufficient flexibility for organizations to prioritize their investments and activities to balance competing resource demands. SMBs can still consult the best practices, standards, and guidelines shared within this report for incremental, practical actions to improve their SCRM posture. A small business should strive to identify risks associated with its unique supply chain partners, and then mitigate the risks, to the best of its ability, consistent with the organization’s risk tolerance.

## Supply Chain Critical Infrastructure

The Information and Communications Technology (ICT) supply chain infrastructure is the integrated set of components (hardware, software and processes) within the organizational boundary that composes the environment in which a system is developed or manufactured, tested, deployed, maintained, and retired/decommissioned.

According to Title 42 US Code, Section 5195c(e), critical infrastructure is defined as the “systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters”. The Department of Homeland Security is currently designated as the Sector-Specific Agency for protection of critical infrastructure vulnerabilities for the ICT sector.[[35]](#footnote-35)

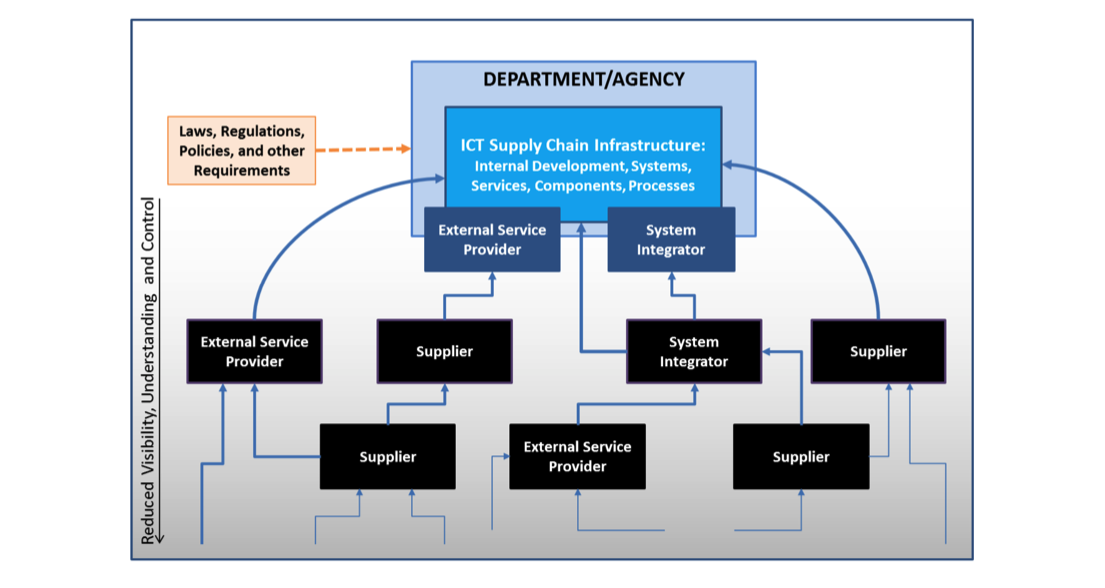


Figure ‑: The high-level model of a supply chain. SOURCE: NIST Special Publication 800-161, Supply Chain Risk Management Practices for Federal Information Systems and Organizations.

Figure 8-1 showcases the high-level model of the supply chain infrastructure, including internal development, information, information systems, government agencies, services, components, and processes to create, maintain, and retire an organization’s information systems.

## Supply Chain Vulnerabilities in 5G

5G is in the early stages of its evolution, as is the corresponding supply chain. A comprehensive risk assessment may be premature. 5G supply chain naturally is reliant on a coherent and healthy global system that leads to risks that occur due to all the moving parts needed to develop such a large-scale structure. There are, however, some risks directly associated with supply chain at a holistic level that should be considered:

### Inadequate Information Sharing across the Supply Chain

The Information Security Forum stated in a report on “Securing the Supply Chain” that “Sharing information with suppliers is essential, yet the wider our ecosystem becomes, the more it increases the risk of that information being compromised.”[[36]](#footnote-36)

### Challenges Specific to Small and Mid-Sized Businesses

The challenges discussed above are typically compounded for small and medium-sized business, which generally have less control and leverage over their supply chain and fewer resources at their disposal to invest in, for example, cyber data and information sharing. SMBs generally do not have market-based leverage to ensure that their suppliers will even share their security practices – let alone the ability to influence those practices.

### Visibility Risk

“Some of the biggest and most complex supply chains have so many external partners that they are unable to assess the risk of doing business with each one. Many companies aren't fully aware of the scope and seriousness of the issue. They suffer from a "black hole" of undefined information risk, especially when it comes to the extended supply chain,” says Information Security Forum (ISF) chief executive officer Michael de Crespigny." They understand and manage this risk internally, he adds, "but have difficulty identifying and managing [it] across their hundreds of thousands of suppliers.”[[37]](#footnote-37)

## Threat Mitigation

The examples above highlighted the difficulty of anticipating supply chain risks. While there are usually three aspects to mitigation centered on education, upgrading technology, and risk management – the supply chain is essentially an issue of trust because the process is external to all stakeholders vested in the process.

The largest aspect in mitigation to supply chain risks in 5G centers around the establishment of trust. It is very critical that an inclusive and broad approach is made to allow all stakeholders within the system to make decisions and identify their role within the system, as well as further map their security-based dependencies and vulnerabilities placed on other members of the structure.

To mitigate all factors that hamper the trust network is very complex because communication supply chains have so many external partners, they may be unable to assess the risk of doing business with each one. However, several steps can be taken to mitigate factors to ensure risk management is an integral part of the supply chain governance;[[38]](#footnote-38)

### Business Management

Most business strategies and processes relating to supply chain may suffer from inefficiencies because of possible gaps in process, communication, uniformity, and information. This may lead to a lack of transparency, understanding and control over supply and decisions involved in the development, acquisition, and delivery of the products and service itself. Suppliers are often disconnected from the decision-making process and are therefore placed in no “direct control” for an organization. At a business level, the transparency and visibility around supply chain is also very difficult to navigate because the process in itself is external to the organization and has a multitude of moving parts that are not placed in monitored control. This pain point can often be the weakest link in a cyber security management program, and examples exist to validate the risk.

The challenges discussed above are typically compounded for small and medium-sized business, which generally have less control and leverage over their supply chain and fewer resources at their disposal to invest in, for example, cyber data and information sharing.

#### Vendor Risk Management[[39]](#footnote-39)

It is critical that vendor risk is addressed and monitored consistently. Vendor risk management policy is the foundation of any vendor risk management (VRM) program and an area that is often overlooked. Your vendors should agree to whatever is included in your VRM policy.

Managing vendor risk is an ongoing process. Having a VRM policy in place ensures that your organization gets the most risk mitigation benefit from its VRM program in the most efficient manner.

According to ISO 27001:2013 section A.15.1:

“A.15.1.1 Information security policy for supplier relationships – Information security requirements for mitigating the risks associated with supplier’s access to the organization’s assets shall be agreed with the supplier and documented.”

A good practice is to start by setting up a company-wide vendor risk ranking system and categorize each vendor within it. Next, for each rank level determine what you need to do to monitor risk, and how often.

Following are ten steps that many VRM policies across industries should consider for “critical” or “high risk” vendors:

* Identify the supplier and the service/product they provide.
* Depict the process flow through which the supplier provides its product or service.
* Identify the types of information being accessed or touched by the supplier.
* Identify the critical control points in that information flow.
* Identify the controls that should be in place to keep the vendor’s business running and maintain confidentiality, integrity and availability (CIA) of your data while it is in their hands.
* Identify how the vendor will continue to provide services to you during a disaster or outage.
* Identify how the vendor will handle incident management where your company is concerned.
* Establish a main point of contact at the vendor.
* Determine how changes to the above will be handled.
* Determine how often the above steps will be re-verified.

### 5G Supply Chain Threats/Risks:

5G is in the early stages of its evolution, as is the corresponding supply chain. A comprehensive risk assessment may be premature. There are, however, some risks directly associated with supply chain at a holistic level that should be considered. 5G supply chain naturally is reliant on a coherent and healthy global system that leads to risks that occur due to all the moving parts needed to develop such a large-scale structure.

5G supply chain vulnerabilities may be found in:[[40]](#footnote-40)

* 1. The systems/components within the software development life cycle (SDLC) (i.e., being developed and integrated);
  2. The development and operational environment directly impacting the SDLC; and
  3. The logistics/delivery environment that transports ICT systems and components (logically or physically)

Use cases, SMEs, and further analyses have led to seven vulnerable elements leading to supply chain risks in 5G;

* Core infrastructure
* Devices
* Radio networks
* Information Technology Platforms
* Cloud Service Providers
* Application Providers and Developers
* Middleware (IoT, Virtualization, etc.)

| Vulnerability Considerations | Methods |
| --- | --- |
| * Organization’s mission/business * Supplier relationships (e.g., system integrators, COTS, external services) * Geographical considerations related to the extent of the organization’s ICT supply chain * Enterprise/Security Architecture * Criticality Baseline | * Examine agency ICT supply chain information including that from supply chain maps to identify especially vulnerable locations or organizations. * Analyze agency mission for susceptibility to potential supply chain vulnerabilities. * Examine system integrator and supplier relationships for susceptibility to potential supply chain vulnerabilities. * Review enterprise architecture and criticality baseline to identify areas of weakness requiring more robust ICT supply chain considerations. |
| * Mission functions * Geographic locations * Types of suppliers (COTS, custom, etc.) * Technologies used | * Refine analysis from the above based on specific mission functions and applicable threat and supply chain information. * Consider using the NVD, including CVE and Common Vulnerability Scoring System (CVSS), to characterize, categorize, and score vulnerabilities. * Consider using scoring guidance to prioritize vulnerabilities for remediation. |
| * Individual technologies, solutions, and suppliers should be considered. | * Use CVEs where available to characterize and categorize vulnerabilities. * Identify weaknesses. |

Table ‑: Potential vulnerabilities and mitigation. SOURCE: NIST SP 800-161

#### Supply Chain Threat Agents in 5G

A supply chain’s global landscape also poses unique risks to the system. The supply chain ecosystem is not simply disrupted via physical and regulatory conditions, but a *convergence* of global actors and risks seeking to maximize personal outcomes. Threat agents are consistently changing, adapting, and growing across a global landscape, and may be external to the supply chain stakeholders themselves. A complex supply chain can often become a playground for nefarious actors with adaptive methods, who may also potentially set conditions for newer actors and methods to exploit vulnerabilities in the ecosystem. Another problem with the globally extended supply chain is the lack of visibility across the landscape. Supply chains are managed in discrete segments, which may operate independently and in isolation. Therefore, a common practice across the system focusing on global end-to-end security and the maintenance of trust is very important but should be catered and framed independently for each system.

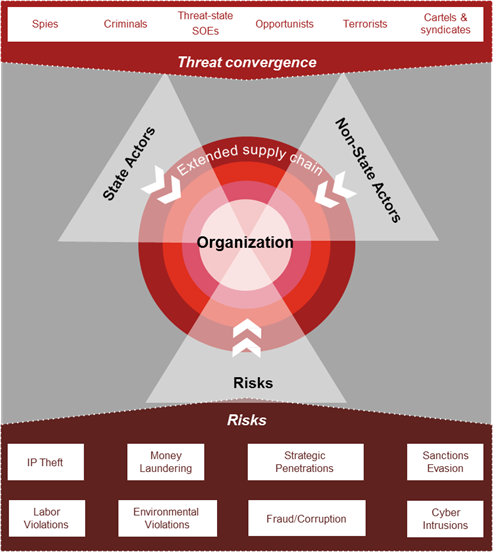


Figure ‑: “Supply Chain Threat Convergence- A New Reality for Supply Chains”. SOURCE: PwC

PwC presented the figure above showcasing a visual of the “Global Threat Convergence” for supply chain, noting that each actor and risk is consistently changing shape as their objectives, methods, and means are concurrently shifting. With this convergence, Industry consistently has more ambiguity to the nature of threat in supply chain due to the complexity and expansiveness of threat actors working to exploit the system. As different threat actors are driven by different motives and with differing access to resources, it is imperative to consistently apply appropriate SCRM procedures across the enterprise supply chain with a focus on the mitigation of risk as opposed to specific actors.

Table 8-2 below, from NIST SP 800-161, lists examples of non-state threat agents and exploitations that are prevalent to the supply chain.

| Threat Agent | Scenario | Examples |
| --- | --- | --- |
| Counterfeiters | Counterfeits inserted into ICT supply chain | Criminal groups seek to acquire and sell counterfeit ICT components for monetary gain. Specifically, organized crime groups seek disposed units, purchase overstock items, and acquire blueprints to obtain ICT components that they can sell through various gray market resellers to acquirers. |
| Insiders | Intellectual property loss or code insertion | Disgruntled insiders sell or transfer intellectual property to competitors or foreign intelligence agencies for a variety of reasons including monetary gain. Intellectual property includes software code, blueprints, or documentation |
| Foreign Intelligence services/Nation States | Malicious code insertion | Foreign intelligence and/or nation states seek to penetrate ICT supply chain and implant unwanted functionality to be used when the system is operational to gather information or subvert system or mission operations |
| Terrorists | Unauthorized access | Terrorists seek to penetrate or disrupt the ICT supply chain and may implant unwanted functionality to obtain information or cause physical disablement and destruction through ICT |
| Industrial Espionage/Cyber Criminals | Industrial Espionage/Intellectual Property Loss | Industrial spies/cyber criminals seek ways to penetrate ICT supply chain to gather information or subvert system or mission operations (e.g., exploitation of an HVAC contractor to steal credit card information). |

Table ‑: Examples of possible threats. SOURCE: NIST SP 800-161

## Supply Chain Risk Management (SCRM):

The examples above highlighted the difficulty of anticipating supply chain risks. While there are usually three aspects to mitigation centered on education, upgrading technology, and risk management – the supply chain is essentially an in issue of trust because the process is external to all stakeholders vested in the process.

The largest aspect in mitigation to supply chain risks in 5G centers around the establishment of trust. It is very critical that an inclusive and broad approach is made to allow all stakeholders within the system to make decisions and identify their role within the system, as well as further map their security-based dependencies and vulnerabilities placed on other members of the structure.

To mitigate all factors that hamper the trust network is very complex because communication supply chains have so many external partners, they may be unable to assess the risk of doing business with each one. However, several steps can be taken to mitigate factors to ensure risk management is an integral part of the supply chain governance;[[41]](#footnote-41)

The Department of Defense has defined SCRM as,

“A systematic process for managing supply chain risk by identifying susceptibilities, vulnerabilities and threats throughout a “supply chain” and developing mitigation strategies to combat those threats whether presented by the supplier, the supplied product and its subcomponents, or the supply chain.”

A primary objective of cyber SCRM is to identify, assess, and mitigate “products and services that may contain potentially malicious functionality, are counterfeit, or are vulnerable due to poor manufacturing and development practices within the cyber supply chain.”[[42]](#footnote-42) Cyber SCRM activities may include:

* Determining cybersecurity requirements for suppliers,
* Enacting cybersecurity requirements through formal agreement (e.g., contracts),
* Communicating to suppliers how those cybersecurity requirements will be verified and validated,
* Verifying that cybersecurity requirements are met through a variety of assessment methodologies, and
* Governing and managing the above activities.[[43]](#footnote-43)

The NIST is responsible for developing standards, guidelines, tests, and metrics for the protection of non-national security federal information and communication infrastructure. Over the past several years, NIST has collaborated with public and private sector stakeholders to research and develop ICT SCRM tools and metrics, as well as guidelines on mitigation strategies and implementation methodologies.[[44]](#footnote-44)

Risk management is a comprehensive process that requires organizations to:

1. Frame risk (i.e., establish the context for risk-based decisions);
2. Assess risk;
3. Respond to risk once determined; and
4. Monitor risk on an ongoing basis using effective organizational communications and a feedback loop for continuous improvement in the risk-related activities of organizations. Figure 2-3 depicts interrelationships among the risk-management process steps, including the order in which each analysis may be executed and the interactions required to ensure that the analysis is inclusive of the various inputs at the organization, mission, and operations levels.

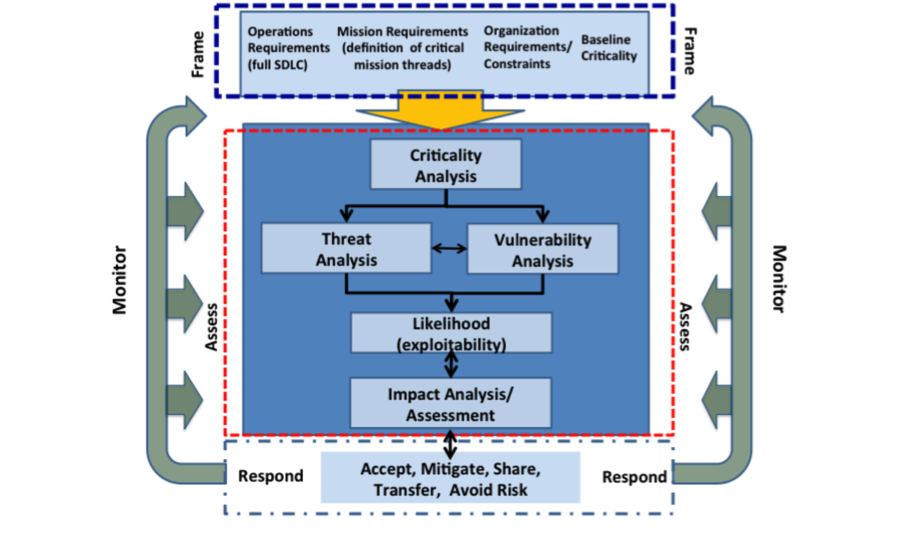


Figure ‑: SCRM activities. SOURCE: NIST SP 800-161

The figure above summarizes SCRM activities throughout the risk management process as they are performed within the three organizational tiers. The arrows between different steps of the risk-management process depict simultaneous flow of information and guidance among the steps. Together the arrows indicate that the inputs, activities, and outputs are continuously interacting and influencing one another. More details are provided in the following subsections.

### NIST Information Communication Technology Risk Management Approach:

Managing ICT supply chain risk requires ensuring the integrity, security, and resilience of the supply chain and its products and services, with their quality also being ensured. NIST has developed an approach to ICT SCRM that encompasses the following key points:

* **Foundational Practices:** ICT SCRM lies at the intersection of information security and supply chain management. Existing supply chain and cybersecurity practices provide a foundation for building an effective ICT SCRM program.
* **Organization-wide:** Effective ICT SCRM is an organization-wide activity that involves each organizational tier (Organization, Mission/Business Processes, and Information Systems) and is implemented throughout the system development life cycle.

**Risk Management Process:** ICT SCRM should be implemented as part of overall risk- management activities, such as those described in NIST SP 800-39, Managing Information Security Risk. Activities should involve identifying and assessing applicable risks, determining appropriate mitigating actions, developing an ICT SCRM Plan to document selected mitigating actions, and monitoring performance against that plan. Because ICT supply chains differ across and within organizations, the ICT SCRM plan should be tailored to individual organizational contexts.

* **Risk:** ICT supply chain risk is associated with a lack of visibility into, understanding of, and control over many of the processes and decisions involved in the development and delivery of ICT products and services acquired by federal agencies.
* **Threats and Vulnerabilities**: Effectively managing ICT supply chain risks requires a comprehensive view of threats and vulnerabilities. Threats can be either “adversarial” (e.g. tampering, counterfeits) or “non-adversarial” (e.g. poor quality, natural disasters); vulnerabilities may be “internal” (e.g. organizational procedures) or “external” (e.g. part of an organization’s supply chain).
* **Critical Systems:** Cost-effective supply chain risk mitigation requires agencies to identify those systems/components that are most vulnerable and will cause the largest organizational impact if compromised.

### 4 Pillars of Supply Chain Risk Management: (Derived from ISO SP 800-61)



Figure ‑: The four pillars of SCRM. SOURCE: ISO SP 800-61.

SCRM holistically encompasses activities in the system development lifecycle, including research and development, design, manufacturing, acquisition, delivery, integration, operations, and disposal/retirement of an organization’s products and services. As ISO 800 states, communication technology SCRM relies on the intersection of the “4 pillars” of SCRM;[[45]](#footnote-45)

1. Security:

Security provides the confidentiality, integrity, and availability of information that (a) describes the ICT supply chain (e.g., information about the paths of ICT products and services, both logical and physical); or (b) traverses the ICT supply chain (e.g., intellectual property contained in ICT products and services), as well as information about the parties participating in the ICT supply chain (anyone who touches an ICT product or service throughout its life cycle).

1. Integrity:

Integrity is focused on ensuring that the ICT products or services in the ICT supply chain are genuine, unaltered, and that the ICT products and services will perform according to acquirer specifications and without additional unwanted functionality.

1. Resilience:

Resilience is focused on ensuring that ICT supply chain will provide required ICT products and services under stress or failure.

1. Quality:

Quality is focused on reducing vulnerabilities that may limit the intended function of a component, lead to component failure, or provide opportunities for exploitation.

### Business Management

Most business strategies and processes relating to supply chain may suffer from inefficiencies because of possible gaps in process, communication, uniformity, and information. This may lead to a lack of transparency, understanding and control over supply and decisions involved in the development, acquisition, and delivery of the products and service itself. Suppliers are often disconnected from the decision-making process and are therefore placed in no “direct control” for an organization. At a business level, the transparency and visibility around supply chain is also very difficult to navigate because the process in itself is external to the organization and has a multitude of moving parts that are not placed in monitored control. This pain point can often be the weakest link in a cyber security management program, and examples exist to validate the risk.

The challenges discussed above are typically compounded for small and medium-sized business, which generally have less control and leverage over their supply chain and fewer resources at their disposal to invest in, for example, cyber data and information sharing.

### Vendor Risk Management[[46]](#footnote-46)

It is critical that vendor risk is addressed and monitored consistently. Vendor risk management policy is the foundation of any vendor risk management (VRM) program and an area that is often overlooked. Your vendors should agree to whatever is included in your VRM policy.

Managing vendor risk is an ongoing process. Having a VRM policy in place ensures that your organization gets the most risk mitigation benefit from its VRM program in the most efficient manner.

According to ISO 27001:2013 section A.15.1:

“A.15.1.1 Information security policy for supplier relationships – Information security requirements for mitigating the risks associated with supplier’s access to the organization’s assets shall be agreed with the supplier and documented.”

A good practice is to start by setting up a company-wide vendor risk ranking system and categorize each vendor within it. Next, for each rank level determine what you need to do to monitor risk, and how often.

Following are ten of the basic steps that many VRM policies across industries should define for “critical” or “high risk” vendors:

* Identify the supplier and the service/product they provide.
* Depict the process flow through which the supplier provides its product or service.
* Identify the types of information being accessed or touched by the supplier.
* Identify the critical control points in that information flow.
* Identify the controls that should be in place to keep the vendor’s business running and maintain the CIA of your data while it is in their hands.
* Identify how the vendor will continue to provide services to you during a disaster or outage.
* Identify how the vendor will handle incident management where your company is concerned.
* Establish a main point of contact at the vendor.
* Determine how changes to the above will be handled.
* Determine how often the above steps will be re-verified.

### Considerations for Small and Mid-Sized Businesses (SMBs)

While the best practices described above are useful in providing a common vocabulary and in describing the manner in which organizations can identify and assess their overall SCRM posture, not all of the guidelines and practices are suitable for all organizations. Perhaps inadvertently, many of these resources imply a level of technical and financial sophistication that may be difficult for SMBs to achieve.

Indeed, smaller communications network operators often have few options for suppliers that meet their needs. In addition, as a byproduct of their size, SMBs are often resource constrained, with limited access to the operational manpower, technical expertise, and financial support required to engage in holistic, company-wide SCRM activities. For instance, SMBs may be unable to enact, evaluate, verify, and validate cybersecurity requirements for their vendors, and they may not have the resources to standup a governance structure allocated to cyber supply chain risk.Perhaps most importantly, SMBs have limited leverage in the marketplace to force suppliers to make wholesale changes to their security practices.

Nevertheless, because the guidance referenced in this report is intended to be tailored to an individual organization’s unique needs, SMBs can still consult the best practices, standards, and guidelines shared within this report for incremental, practical actions to improve their SCRM posture. Small businesses should strive to identify risks associated with their unique supply chain partners, and then mitigate the risks, to the best of their ability, consistent with the organization’s risk tolerance.

Risk management is an integral part of supply chain governance. A number of other steps can be taken to ensure risk management is an explicit and integral part of supply chain governance. Arguably, the initial step may be to institutionalize a multi-stakeholder supply chain risk assessment process that engages as many members of the supply chain as possible. Making sure all stakeholders are involved in this process and ensuring effective information sharing are essential for deterrence of ticketed risk(s) in the system.

It is also important to develop a process so that the risk management steps are taken into account and organization partners are held accountable via contract obligations and remedies, and effective monitoring and enforcement: [[47]](#footnote-47)

* Adhere to procurement processes, evaluating cyber risk from the start
* Conduct due diligence for new suppliers, accounting for their cyber-security competence
* Consider contractual clauses focused on security, stipulating responsibility for any compromise or data breaches and contractually mandate the flow down of security clauses to sub-contractor(s) in supply chain
* Challenge suppliers to practice and develop collaborative processes for reacting to compromise or data breaches
* Conduct regular information assurance activity identifying critical pathways.

### Consideration of Distributed Ledger Technology (DLT) in Supply Chain:

Communication supply chains can span over hundreds of stages, multiple geographic locations, invoices, have several individuals and entities involved, and extend over months of time. Due to the complexity and lack of transparency in current supply chains, there is considerable interest in the application of DLT ([blockchain](https://www.bernardmarr.com/default.asp?contentID=1389) technology) to transform the supply chain and logistics industry.

DLT enables proof of ownership and the transfer of ownership form one entity to another without requiring a third-party intermediary. There is not one central authority over the blockchain, making the system extremely efficient, scalable, and reducing risk as by distributing information across multiple Ledgers as opposed to a single authority.

DLT can be used for any exchange, agreements/contracts, tracking, traceability, payment, etc. Since every transaction is recorded on a block and across multiple copies of the ledger that are distributed over many nodes (computers), each block is recorded, irreversible, and immutable, providing provenance in the data recorded. Multiple players/actors in a supply chain benefit from DLT in that all of them have common Ids and view of the supply chain, instead of dealing with the logistics individually and peace-meal.

DLT is also highly secure since every block links to the one before it and after it. A typical blockchain, for example Hyperledger, comes with a business process layer which can be leveraged for automation and security of the process.

Smart contracts are self-executing contracts with the terms of the agreement between buyer and seller being directly written into lines of code. The code and the agreements contained therein exist across a distributed, decentralized [blockchain](https://www.investopedia.com/terms/b/blockchain.asp) network. The utilization of smart contracts in a supply chain offer opportunity for automation of transactions and engagement of all participants.

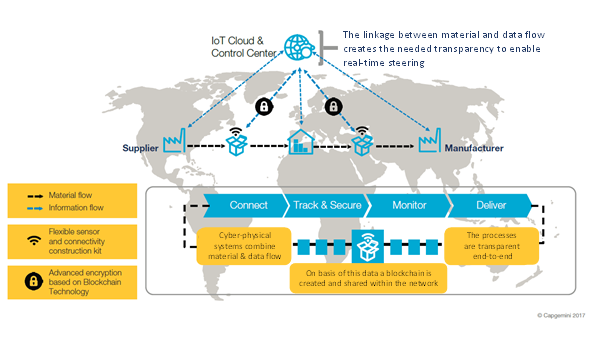


Figure ‑: SCRM and DLT. SOURCE: https://www.capgemini.com/

Figure 8-5 above showcases the visual Capgemini and IBM developed in 2017 as a prototype for Supply Chain Risk Management based on Blockchain technology. The collected data is visible in real-time to all affected members of the Supply Chain through this prototype platform which guarantees data availability and manipulation security.

## Supply Chain Industry Standards:

The following sections highlight examples of the work done by industry, sometimes in partnership with government, to develop standards and frameworks for identifying, assessing and mitigating supply chain risks.

### NIST Cyber Security Framework 1.1 Guidance on Supply Chain Risk Management

The latest draft of the NIST cyber security framework addresses SCRM. Communication is essential to SCRM, and the framework provides common language to communicate requirements among interdependent stakeholders responsible for the delivery of essential critical infrastructure products and services.

The Framework recommends having a “Target Profile” to develop an explicit scaling process for external partners. This is essential in identifying a cohesive SCRM process that prioritizes suppliers based on an objective assessment of risk.

The framework also emphasizes the need for developing an enforcement strategy when utilizing external partners and recommends the following actions:

* Contracts should set obligations to implement security measures, and should be enforced
* Periodic supplier audits
* Suppliers participate in response and recovery planning

The framework also highlights the idea of developing a tier selection process after the target profile has been finalized. The tier selection process considers an organization’s current risk management practices, threat environment, legal and regulatory requirements, information sharing practices, business/mission objectives, supply chain cybersecurity requirements, and organizational constraints. Organizations should determine the desired tier(s) for their suppliers, ensuring that the selected level meets the organizational goals, is feasible to implement, and reduces cybersecurity risk to critical assets and resources to levels acceptable to the organization. Organizations should consider leveraging external guidance obtained from Federal government departments and agencies, Information Sharing and Analysis Centers (ISACs), Information Sharing and Analysis Organizations (ISAOs), existing maturity models, and other sources to assist in determining their desired tier.

| **Category** | **Subcategory (Tiers)** | **Informative References** |
| --- | --- | --- |
| **Supply Chain Risk Management (ID.SC):**  The organization’s priorities, constraints, risk tolerances, and assumptions are established and used to support risk decisions associated with managing supply chain risk. The organization has established and implemented the processes to identify, assess and manage supply chain risks. | **ID.SC-1:** Cyber supply chain risk management processes are identified, established, assessed, managed, and agreed to by organizational stakeholders | **CIS CSC** 4  **COBIT 5** APO10.01, APO10.04, APO12.04, APO12.05, APO13.02, BAI01.03, BAI02.03, BAI04.02  **ISA 62443-2-1:2009** 4.3.4.2  **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3, A.15.2.1, A.15.2.2  **NIST SP 800-53 Rev. 4** SA-9, SA-12, PM-9 |
| **ID.SC-2:** Suppliers and third party partners of information systems, components, and services are identified, prioritized, and assessed using a cyber supply chain risk assessment process | **COBIT 5** APO10.01, APO10.02, APO10.04, APO10.05, APO12.01, APO12.02, APO12.03, APO12.04, APO12.05, APO12.06, APO13.02, BAI02.03  **ISA 62443-2-1:2009** 4.2.3.1, 4.2.3.2, 4.2.3.3, 4.2.3.4, 4.2.3.6, 4.2.3.8, 4.2.3.9, 4.2.3.10, 4.2.3.12, 4.2.3.13, 4.2.3.14  **ISO/IEC 27001:2013** A.15.2.1, A.15.2.2  **NIST SP 800-53 Rev. 4** RA-2, RA-3, SA-12, SA14, SA-15, PM-9 |
| **ID.SC-3:** Contracts with suppliers and third-party partners are used to implement appropriate measures designed to meet the objectives of an organization’s cybersecurity program and Cyber Supply Chain Risk Management Plan. | **COBIT 5** APO10.01, APO10.02, APO10.03, APO10.04, APO10.05  **ISA 62443-2-1:2009** 4.3.2.6.4, 4.3.2.6.7  **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3  **NIST SP 800-53 Rev. 4** SA-9, SA-11, SA-12, PM9 |
| **ID.SC-4:** Suppliers and third-party partners are routinely assessed using audits, test results, or other forms of evaluations to confirm they are meeting their contractual obligations. | **COBIT 5** APO10.01, APO10.03, APO10.04, APO10.05, MEA01.01, MEA01.02, MEA01.03, MEA01.04, MEA01.05  **ISA 62443-2-1:2009** 4.3.2.6.7  **ISA 62443-3-3:2013** SR 6.1  **ISO/IEC 27001:2013** A.15.2.1, A.15.2.2  **NIST SP 800-53 Rev. 4** AU-2, AU-6, AU-12, AU16, PS-7, SA-9, SA-12 |

Table ‑:Risk management standards. SOURCE: NIST.

## NIST SP 800-161; Supply Chain Risk Management Practices for Federal Information Systems and Organizations:

NIST SP 800-161 provides federal agencies with guidance to develop the appropriate policies, processes, and controls to effectively manage ICT supply chain risk. It is flexible and builds on agencies’ existing information security practices.

* **Risk Management:** NIST SP 800-161 details a set of processes for evaluating and managing supply chain risk. These processes are integrated into the NIST SP 800-39’s Risk Management Process (Frame, Assess, Respond, and Monitor) and should be implemented as part of agencies’ overall risk management activities.
* **Extended Overlay:** Several controls in Appendix F of NIST SP 800-53 Rev. 4 can help with ICT supply chain risk mitigation. Chapter 3 of NIST SP 800-161 identifies these controls and provides supplementary guidance for their application to ICT SCRM. Additional controls assist organizations in developing more robust and complete ICT SCRM mitigation strategies.
* **Threat Scenarios and Risk Framework:** Understanding and evaluating ICT SCRM threats supports a cost-effective risk mitigation strategy. NIST SP 800-161 lists applicable threat events and provides a risk framework for assessing threats and identifying mitigation responses—one method for evaluating interdependencies and the potential impact of an event.
* **ICT SCRM Plan:** NIST SP 800-161 provides a template for developing ICT SCRM plans that address the entire system development life cycle.

### International Organization for Standardization Supply Chain Standards:

The International Organization for Standardization (ISO) produced the ISO 27000 and 28000 series of standards, specifically written to address ICT security matters. ISO 27001 addresses process and auditing standards, which takes companies from basic risk assessments through to policies for managing information, communications, human resources, physical sites, business continuity and compliance. ISO 28000 series address security management for supply chains.

Adopting the ISO 31000 series (addressing risk management) supplementary to the 28000 and 27000 then provides a common language to communicate cyber security risk used throughout the supply chain. The importance of the international standards is in their use by all stakeholders in the same line of business. If all supply chain ecosystem members are utilizing and/or certified to the standards then they share a common understanding of risk, and have a baseline set of measures in place for management therefore making all involved more resilient to the risk.

### ISO 27000 series; Information Security Standards[[48]](#footnote-48)

ISO 27000 series addresses supply chain risk management and information security management. ISO 27036 noted that a common problem is that the acquirer and supplier are not sharing equal responsibility for making sure their agreement is trustworthy and for managing their information security risks that can include delineated roles and responsibilities around information security. To manage security risks and vulnerabilities the acquirer should control the supplier’s access to information. ISO 27001 formally specifies how to establish an Information Security Management System (ISMS). As NIST highlights, ISO 270001 states that it is critical we define an ISMS scope. Management of the industry should define the scope of the ISMS in terms of the nature of the business, the organization, its location, information assets and technologies. Any exclusion from the ISMS scope should be justified and documented. Each supplier relationship is established for a specific purpose. The number of such relationships is likely to grow over time, increasing the complexity of managing the entire supplier base. ISO 27036 recommends that when acquiring services the acquirer should establish rules for what information is critical to share with the supplier, and how to control sharing of information.

### ISO 28000 Series; Supply Chain Security Management:

Supply chain security management systems based on the ISO 28000 certification standard identify the risk levels across your supply chain operations. This information then enables your organization to carry out risk assessments and apply the necessary controls with supporting management tools (i.e. document controls, key performance indicators, internal audit and training).[[49]](#footnote-49) The requirements for ISO 28000 include all critical aspects for supply chain security assurance. Some examples include: financing, manufacturing, information management, and the facilities for packing, storing and transferring goods between vehicles and locations. Security management is linked to many other aspects of business management. These should be considered directly, when and where they impact security management, including transporting goods through the supply chain. ISO 28000 provides a framework (figure below) for providing effective physical security management through a system that identifies security threats, assesses risk, establishes objectives for implementing controls and continuously improves the physical security of the organization.[[50]](#footnote-50)

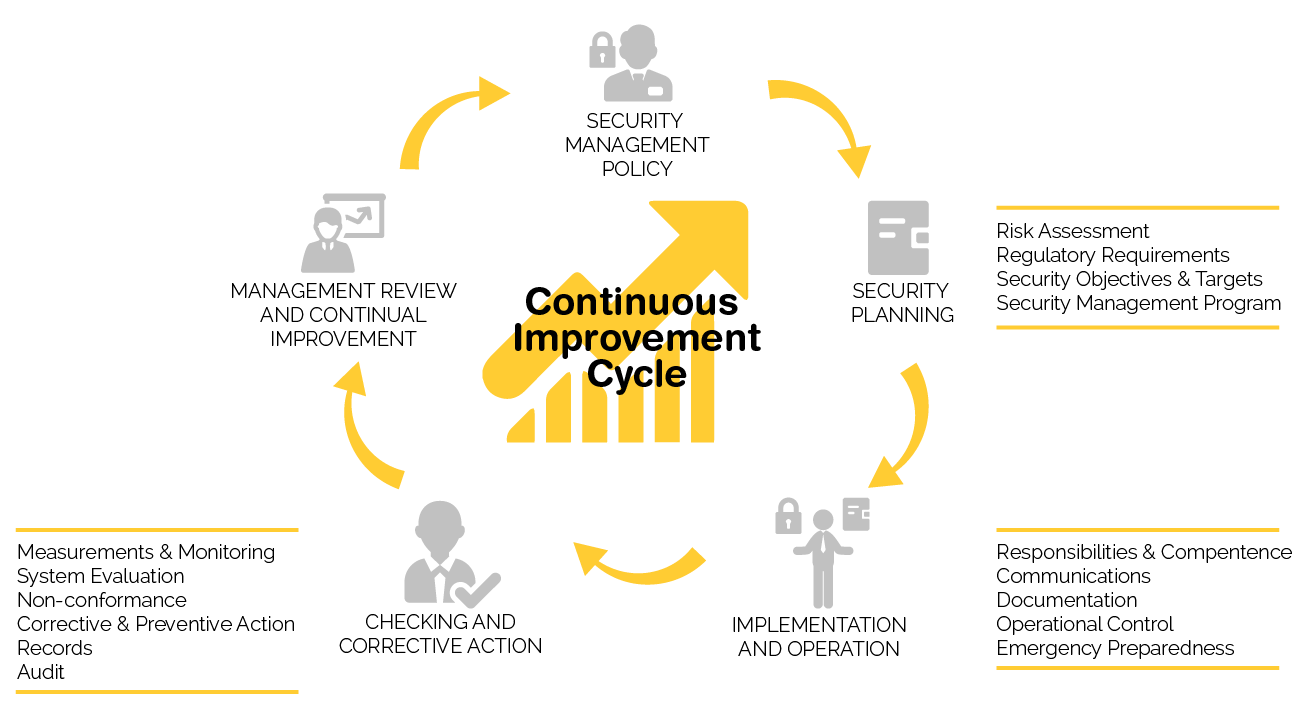


Figure ‑: ISO 28000 SCRM Management Framework. SOURCE: ISO

The 28000 series consists of the following standards:

* ISO 28000:2007, Specification for security management systems for the supply chain;
* ISO 28001:2007, Security management systems for the supply chain – Best practices for implementing supply chain security – Assessments and plans – Requirements and guidance;
* ISO 28003:2007, Security management systems for the supply chain – Requirements for bodies providing audit and certification of supply chain security management systems;
* ISO 28004:2007, Security management systems for the supply chain – Guidelines for the implementation of ISO 28000.

# Conclusions

## Topic of Report

The telecommunications industry engaged in the evolution of wireless networks to a completely new technology. 5G is the commonly used acronym assigned to the 5th generation of wireless technology that has both a fixed and a mobile component and is both terrestrial and non-terrestrial based. It is much broader than previous generations of wireless networks. While previous generations of network services were focused on a single type of network (i.e., mobile wireless), 5G has been defined as a “network of networks.” This is because of the recognition that a single technology (e.g., mobile wireless) will not be able to provide all the characteristics of 5G such as ubiquity and resiliency.

The CSRIC VI, WG3 has focused on four main areas of 5G technology; IoT, NFV/SDN, open source software, and supply chain risks. Maintaining the focus on 5G was critical to meet the stringent deadline for this report. While there are vulnerabilities in these areas in 3G and 4G, the charter of this working group was to study the risks and vulnerabilities in 5G networks only.

This report begins with a description of 5G that highlights some of the key areas of technology that will change in the radio and core network. The data gathering and analysis to support our findings address the core network technology of 5G: network security specifications from 3GPP, IoT connectivity, and NFV/SDN. In addition, the report examines the threats and benefits introduced by the use of open source software, which is found even today in networks.

## Method of Data Gathering

The CSRIC VI WG3 was asked to examine the security risks associated with 5G that result from both infrastructure and device changes that may introduce incremental security risk. Specifically, the Working Group was tasked to evaluate security risks in four domains stemming from the evolution to 5G:

* IoT
* OSS
* NFV and SDN

The 3GPP, as well as several other standards organizations, continues work on 5G standards. A full 5G network core will not be likely until after 2020, and even then, on a limited basis. Wide scale deployment of 5G network core will take time to implement, as we have seen with other technologies (including 4G). As a result, the body of work on threats, risks and best practices to mitigate risk to 5G networks is still maturing. To address this limitation, Working Group 3 drew upon several sources to compile the data to identify and evaluate the emerging security risks anticipated for 5G:

* Industry SME presentations
* Standards bodies and industry associations (GSMA, CTIA, ETSI, 3GPP, NIST, ISO)
* Individual contributor research gathered by Working Group members

Some of the specific topics under consideration by WG3 arise because of the migration away from traditional, engineered systems designed to support specific network functions to a more distributed, software-based architecture. This new architecture exposes telecom networks to new attack vectors stemming from the adoption of IT technologies. As a result, the research into OSS, IoT, and even NFV/SDN (to a lesser extent) are influenced by body of work addressing risk mitigation in the IT domain where these capabilities have existed for a number of years.

## Approach to Analysis

As 5G standards continue to mature, and operators begin planning and executing on their deployments of 5G networks, security is top-of-mind. The risks identified in prior technologies are still applicable to 5G. 5G will also introduce new attack vectors and a completely new network architecture that is software-based and embraces virtualization. IoT, SDNs, and open source software will be a critical ingredient in the 5G ecosystem. This expanded threat environment is being addressed by industry up front through the incorporation of security improvements early in the development of the architecture and specifications for 5G.

The CSRIC WG3 was tasked to evaluate mechanisms to best design and deploy 5G networks to mitigate risks to network reliability and security posed by the evolution to a software-based SOA, the proliferation of IoT devices and expanding use of open source platforms used in 5G networks.

## Basis of Recommendations

As part of this assessment, WG3 developed risk impacts for each of the key areas and provided recommendations to mitigate the identified risks and best practices within design, deployment and operation of risk-tolerant 5G. Analysis of the data gathered concerning risks and approaches to mitigation are presented with respect to the each of these technologies. The resulting best practices are captured and documented as recommendations for industry and the FCC to be considered by CSRIC VI.

The 5G standards address a number of vulnerabilities that have impacted previous network evolutions. 5G will be more resilient and introduces new security procedures, however 5G does introduce new attack vectors and requires a solid security framework to be implemented throughout the network. Fortunately, security considerations are an integral design consideration in the development of the 5G architecture specifications and standards that will result in a more resilient and secure framework for 5G network implementation.

# 

# Final CSRIC Recommendations for the FCC

## General Recommendations

The CSRIC makes the following general recommendations for the FCC:

### Extension of CSRIC 5G Study

CSRIC recommends that the FCC extend the CSRIC 5G research in the next CSRIC, given the immaturity of 5G technology at this time. This CSRIC working group believes that there will be new findings and conclusions once the 3GPP (and other) standards on 5G and associated technologies are completed (which may be another year or more), and once 5G deployments begin.

Given the current state of the standards and deployment, the working group therefore feels it would be best to revisit 5G in the next CSRIC.

### Threat Taxonomy

FCC should work with industry, DHS and NIST through a future CSRIC or other collaborative process to develop a communications sector-specific taxonomy for communicating threats and evaluating risks to the 5G ecosystem. The focus of this proposed work should be to provide the best available information and tools to enable industry and government to make more informed risk management decisions.

### Participate with DHS in Public/Private Partnership for Threat Assessment

With the evolution towards a 5G network architecture that combines communications network technology with established information technologies (i.e. a ‘converged’ network), the need to effectively and quickly understand, assess, and mitigate emerging threats to the ecosystem will become more critical because of the wide variety of systems, applications, and users that depend upon the network. Existing resources are already heavily taxed in responding to the steady stream of alert notifications received from multiple distinct threat information source feeds; the key challenges lie in identifying the most potentially impactful issues and in prioritizing them. The SMB community in particular would benefit by receiving an accurate assessment of the threat in addition to the notification of its existence. CSRIC WG3 recommends that the FCC partner with DHS and Industry to evaluate options for providing a realistic, timely, and prioritized risk assessment of identified and potential threats.

## Recommendations for IoT

The CSRIC makes the following IoT recommendations for the FCC:

### Regulatory actions

CSRIC recommends that the FCC avoid IoT regulatory actions since there is an ongoing industry wide effort, through innovation and collaboration, to evolve security technologies and controls to mitigate the potential risks facilitated by IoT and 5G.[[51]](#footnote-51)

### Public/Private Partnership

The CSRIC recommends that the FCC should continue to participate with the Departments of Commerce and Homeland Security as they coordinate with industry, civil society, and international partners to develop and execute the roadmap of prioritized actions to increase the resilience of the Internet and communications ecosystem against distributed threats.

## Recommendations for NFV/SDN

The CSRIC makes the following NFV/SDN recommendations for the FCC:

### Regulatory Action

Suppliers and operators are developing and deploying networks that encompass NFV/SDN. Therefore, the inherent security advantages associated with NFV/SDN will apply to 5G, and industry innovation and competition will continue to enhance and drive security evolution in this space. For that reason, CSRIC recommends that the FCC avoid NFV/SDN regulatory actions.

## Recommendations for Open Source

The CSRIC makes the following open source recommendations for the FCC:

### Regulatory Action

CSRIC recommends that the FCC avoid open source regulatory actions since operators, suppliers and standards bodies are collaborating on open source software development for platforms that will support 5G; for example, ONAP. This industry collaboration will facilitate security best practice and controls such as:

* Open source code contributors must be verified so that they are accountable for the code that they produce.
* Code management systems, such as digital signing, must be in place to ensure nonrepudiation of the code.
* A rating system for open source code contributors should help increase the credibility of open source software.
* A security framework is needed to handle the new release velocity (continuous release cycles) for open source software that is envisioned.
* Open source code must be rigorously vetted prior to being placed into production.  This vetting process must be defined and continually reviewed and improved.  It should include:
  + Approval by an internal governance board.
  + Ongoing code scanning (both static and dynamic) for security vulnerabilities.
  + Pre-production steps such as platform hardening and penetration testing.

# Final CSRIC Recommendations for Industry

## General Recommendations

The CSRIC makes the following general recommendations for industry:

### Protecting the 4G Core

As many network operators will be launching 5G radio supported by a 4G EPC, it is imperative that network operators take measures to implement the recommendations of CSRIC VI, WG3 Diameter Best Practices.

### REST API Authorization

It is necessary that network operators configure the authorization of the REST APIs correctly for a successful implementation of the end-to-end core network interconnection security solution specified in 3GPP TS 33.501. Only network operators know who their roaming partners are and what information element the partners should be authorized to access. Grouping of roaming contracts into authorization classes will make the configurations less complex.

### REST API Authentication

The right credentials need to be in place to confirm the authenticity of requests. The credentials need to be exchanged with partners (which can be part of the normal message flow) and kept up-to-date with respect to revocation and expiry to avoid business interruptions. Security updates are essential, as REST API is also commonly used in the Internet and exploits are readily available.

### Protection of the User Data

3GPP 33.501 recommends vendor support of ciphering for the confidentiality of user data, and integrity protection of user data between the UE and the gNB, but it is optional for service providers to use.

CSRIC recommends adoption of ciphering for the confidentiality of user data, and integrity protection of user data between the UE and the gNB.

### Protecting RRC and NAS Signaling

3GPP 33.501 recommends vendor support for protection of the RRC-signaling and NAS signaling, but optional for service providers to use.

CSRIC recommends protection of the RRC-signaling and NAS signaling.

## Recommendations for IoT

The CSRIC makes the following industry recommendations for IoT:

### Industry Collaboration

Since IoT networks will be a network of networks in many cases, it is critical that these different network operators work together on an industry-wide basis to minimize any risks. CSRIC recommends that these sectors work together to develop best practices or other ways to best minimize threats to IoT.

### NB-IoT Management

CSRIC recommends that industry implement 3GPP controls for management of NB-IoT traffic to prevent overloading the network.

### IoT Best Practices

CSRIC recommends that industry evaluate for applicability the GSMA CLP.14, API and SOA security best practices to mitigate IoT threats included in this report for voluntary adoption as applicable and if appropriate.

### **CSRIC IV, WG-5 Remediation of Server-Based DDoS Attacks**

The CSRIC recommends that industry follow the recommendations of CSRIC IV, WG5 for the mitigation of DDoS attacks.

### CSRIC III, WG-7 Botnet Remediation and the Anti-Botnet Code of **Conduct**

The CSRIC recommends industry follow the recommendations of CSRIC III, WG7 for the mitigation of botnets.

### **CSRIC II, WG-8 ISP Network Protection Practices**

The CSRIC recommends industry follow the recommendations of CSRIC II, WG8 of best practices for protecting end-users and the network.

### Comm Sector Coordinating Council White Paper

The CSRIC recommends industry collaborate with other service providers, as well as government agencies such as the FCC and DHS, to help neutralize the threats from botnets, as laid out by the CSCC in its white paper created in response to Executive Order 13800.

## Recommendations for NFV/SDN

The CSRIC makes the following industry recommendations for NFV/SDN:

### Monitor, Monitor and Then Analyze

CSRIC recommends monitoring and analytics be implemented. Strong analytics with machine learning and AI may be necessary to detect anomalies between network functions.

### Adopt Industry Driven Standards

CSRIC recommends industry adopt as appropriate the recommendations provided in the ETSI white paper addressing NFV; Ecosystem; Report on SDN Usage in NFV Architectural Framework.[[52]](#footnote-52)

### Future CSRIC Study on NFV/SDN

CSRIC recommends further study addressing NFV management, forwarding plane architecture, and the control network to mitigate DoS exploits on the network.

## Recommendations for Open Source

The CSRIC makes the following industry recommendations for Open Source:

### Inventory Management of Open Source Components

CSRIC recommends vendors develop procedures for inventory of open source components used in internal development efforts. While these procedures will vary according to the specific situation of each company, they will enable vendors to continually monitor for CVEs against any of the open source components that may be used within one of their products. Protection of this information may be critical to prevent system wide vulnerabilities being disclosed.

### Code Management Systems

CSRIC recommends that a code management system must be in place, such as digital signing, to ensure nonrepudiation of the code.

### Testing of Open Source Code

CSRIC recommends that when using open source code for internal development, the code must be rigorously tested in accordance with industry best practices prior to being placed into production.

### Security Frameworks

A framework should be developed for managing the maintenance cycle of all open source components used within a vendor’s products, or within an operator’s network. This framework must be able to support rapid implementations of maintenance patches developed for specific vulnerabilities in open source.

# Appendix A – Acronyms

AF Application Function

AKA Authentication and Key Agreement

AMF Access and Mobility Management Function

AN Access network

API Application programming interface

APN Access Point Name

ARP Access Resolution Protocol

AS Application Server

ATIS Alliance for Telecom Industry Solutions

AUSF Authentication Server Function

AWS Amazon Web Services

BB-IoT Broadband Internet of Things

BGP Border Gateway Protocol

BP Best Practice

CBRS Citizen Band Radio Service

CBSD Citizens Broadband Radio Service Devices

CI Critical Infrastructure

CIA Confidentiality, integrity and availability

CNI Critical national infrastructure

COTS Commercial Off The Shelf

CSCC Communications Sector Coordinating Council

CSRF Cross-Site Request Forgery

CSRIC Communications Security, Reliability and Interoperability Council

CVE Common Vulnerabilities and Exposures

CVSS Common Vulnerability Scoring System

DDoS Distributed Denial of Service

DHS Department of Homeland Security

DLT Distributed Ledger Technology

DNS Domain Name Server

DoC Department of Commerce

DOM Document Object Model

DoS Denial of Service

DRA Diameter Routing Agent

e2e End-to-end

EAP Extensible authentication protocol

EDGE Enhanced Data rates for GSM Evolution

eMBB Enhanced Mobile Broadband

EPC Evolved packet core

FQDN Fully Qualified Domain Names

gNB Next generation node B

GSM Global System for Mobile communication

GUTI Globally Unique Temporary Identifier

HAPS High altitude platforms

HMAC Hash-based Message Authentication Code

HSS Home Subscriber Server

HTTP Hypertext Transfer Protocol

HTTPS Hypertext Transfer Protocol - Secure

IDS Intrusion Detection System

IETF Internet Engineering Task Force

IMEI International Mobile Equipment Identifier

IMS IP Multimedia Subsystem

IMSI International mobile subscriber identity

IoT Internet of things

IP Internet Protocol

IPS Intrusion Prevention Systems

IPX Internetwork Packet Exchange

ISF Information Security Forum

ISG Industry specification group

ISMS Information Security Management System

ISP Internet Service Provider

JSON JavaScript Object Notation

JWT JSON Web Token

KPI Key Performance Indicator

LDPC Low density parity check

LPWA Low Power Wide Area

LTE Long Term Evolution

OWASP Open Web Application Security Project

M2M Machine-to-machine

MAC Message authentication codes

MANO Management and Orchestration

MCC Mobile country code

ME Mobile Equipment

MEC Multi-access edge computing

MIME Multipurpose Internet Mail Extensions

MIMO Multiple input multiple output

MME Mobility Management Entity

mMTC massive Machine Type Communications

MNC Mobile network code

N3IWF Non-3GPP interworking function

NAI Network Access Identifier

NAS Non Access Stratum

NB-IoT Narrow Band Internet of Things

NEF Network Exposure Function

NF Network Function

NFV Network function virtualization

NFVI Network Function Virtualization Infrastructure

NG Next generation

NGMN Next Generation Mobile Networks

NIA New IP Agency

NR New radio

NRF Network Repository Function

NSI Network slice instance

NSP Node Security Project

NSSF Network Slice Selection Function

NVD National Vulnerability Database

OAM Operations, administration, and maintenance

OFDM Orthogonal frequency-division multiplexing

ONAP Open Network Automation Platform

OS Operating System

OSS Open source software

OTA Over the air

OWASP Open Web Application Security Project

PCF Policy Control Function

PCRF Policy and Charging Rules Function

PEI Permanent Equipment Identifier

PGW Packet Data Network Gateway

PLMN Public land mobile network

QoS Quality of Service

RAN Radio access network

REST Representational State Transfer

RRC Radio Resource Control

SAS Spectrum Access System

SBA Service based architecture

SCEF Service Capability Exposure Function

SCRM Supply Chain Risk Management

SCS Service Capability Server

SDLC Software development life cycle

SDN Software defined network

SEAF Security Anchor Function

SEPP Security edge protection proxy

SGW Serving Gateway

SIDF Subscription identifier de-concealing function

SLAs Service level agreements

SMB Small and Mid-Sized Businesses

SME Subject Matter Expert

SMF Session Management Function

SN Serving network

SOA Service Oriented Architecture

SOAP Simple Object Access Protocol

SQL Structured Query Language

SSIG Standardization Special Interest Group

SSL Secure Sockets Layer

STP Signal Transfer Part

SUCI Subscription Concealed Identifier

SUPI Subscriber Permanent Identifier

TCP Transmission Control Protocol

TLS Transport Layer Security

TMSI Temporary Mobile Subscriber Identity

UDM Unified Data Management

UE User equipment

UPF User Plane Function

URI Universal Resource Identifier

URL Universal Resource Locator

URLLC Ultra Reliable and Low Latency Communications

USIM Universal Subscriber Identity Module

V2V Vehicle to Vehicle

VNF Virtual Network Function

VoIP Voice over IP

WCDMA Wideband Code Division Multiple Access

WG3 Working Group 3

XML Extensible Markup Language

XSS Cross-Site Scripting

1. CTIA Whitepaper – Protecting America’s Next Generation Networks https://www.ctia.org/news/protecting-americas-next-generation-networks [↑](#footnote-ref-1)
2. http://www.3gpp.org/about-3gpp [↑](#footnote-ref-2)
3. Fierce Wireless; “How ITU, 5GPPP, NGMN and others will create the standard for 5G” [↑](#footnote-ref-3)
4. <http://www.rfwireless-world.com/Terminology/5G-KPIs-Key-Performance-Indicators.html> [↑](#footnote-ref-4)
5. Cisco 2017 Visual Networking Index forecast (VNI) [↑](#footnote-ref-5)
6. 3GPP uses the term, Service Based Middleware (SBA) for the specific instantiation of Service Oriented Architecture in 5G specifications. Service Oriented Architecture (SOA) is the general purpose and more broadly used nomenclature. WG3 uses SOA in this report as the acronym of choice. [↑](#footnote-ref-6)
7. Throughout this report, the term “Network Operator” refers to companies that own or operate communications infrastructure that enables connectivity. The term “Service Provider” refers to providers of end user services or cloud-based applications. As noted, a Network Operator can also be a Service Provider. [↑](#footnote-ref-7)
8. ETSI Multi-access Edge Computing Industry Specification Group, https://www.etsi.org/technologies-clusters/technologies/multi-access-edge-computing [↑](#footnote-ref-8)
9. 3GPP uses the term, Service Based Middleware (SBA) for the specific instantiation of Service Oriented Architecture in 5G specifications. Service Oriented Architecture (SOA) is the general purpose and more broadly used nomenclature. WG3 uses SOA in this report as the acronym of choice. [↑](#footnote-ref-9)
10. The White House Office of the Press Secretary, *Presidential Executive Order on Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure* (May 11, 2017), available [at https://www.whitehouse.gov/the-press-office/2017/05/11/presidential-executive-order-strengthening-cybersecurity-federal](file:///D:\Users\Suzon.Cameron\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\CM4MIPGN\at%20%20https:\www.whitehouse.gov\the-press-office\2017\05\11\presidential-executive-order-strengthening-cybersecurity-federal). [↑](#footnote-ref-10)
11. *NSTAC Report to the President on Internet and Communications Resilience* (2017) [↑](#footnote-ref-11)
12. 5G Services & Use Cases, November 2017, 5G Americas Whitepaper [↑](#footnote-ref-12)
13. Statista.com, https://www.statista.com/statistics/275849/number-of-vehicles-connected-to-the-internet/ [↑](#footnote-ref-13)
14. IoT Security Guidelines, Overview Document, CLP.11, GSMA Association, Version 2.0, 31 October 2017 [↑](#footnote-ref-14)
15. 3GPP Definition for LTE networks. [↑](#footnote-ref-15)
16. F5, “The Internet of Things: Security and Business Impacts on Service Providers”, White Paper by Patricia Du [↑](#footnote-ref-16)
17. “API Security: A Disjointed Affair”, April 2016, Rik Turner, OVUM [↑](#footnote-ref-17)
18. “Should REST API Security Risks Keep You Awake At Night?”, August 2017, Philipp Schöne [↑](#footnote-ref-18)
19. https://apifriends.com/2017/08/03/rest-api-security/ [↑](#footnote-ref-19)
20. “Top 5 REST API Security Guidelines”, December 2016, Guy Levin [↑](#footnote-ref-20)
21. “REST Security Cheat Sheet”, November 2017, various authors, OWASP.org [↑](#footnote-ref-21)
22. NFV and SDN are complimentary architectures that are often presented together. However, they are not required to be deployed together in a 5G network core. As network operators proceed with 5G networks in phased deployments, it will be common to see NFV deployed first. [↑](#footnote-ref-22)
23. https://blog.risingstack.com/operating-system-containers-vs-application-containers/ [↑](#footnote-ref-23)
24. ETSI GS NFV-SWA 001 V1.1.1 [↑](#footnote-ref-24)
25. For example, see ETSI GS NFV-SEC 013 v3.1.1 (2017-02), *Network Functions Virtualisation (NFV) Release 3; Security; Security Management and Monitoring Specification* [↑](#footnote-ref-25)
26. Oracle White Paper, September 2013, The Department of Defense (DoD) and Open Source Software [↑](#footnote-ref-26)
27. Compiled by Sherif Koussa, CEO, Software Secured and published in TechBeacon; https://techbeacon.com/13-tools-checking-security-risk-open-source-dependencies-0 [↑](#footnote-ref-27)
28. https://www.owasp.org/index.php/OWASP\_Dependency\_Check [↑](#footnote-ref-28)
29. https://github.com/rubysec/bundler-audit [↑](#footnote-ref-29)
30. https://hakiri.io/ [↑](#footnote-ref-30)
31. https://snyk.io/ [↑](#footnote-ref-31)
32. https://gemnasium.com/ [↑](#footnote-ref-32)
33. https://srcclr.com/ [↑](#footnote-ref-33)
34. Internet Security Alliance (ISA) [↑](#footnote-ref-34)
35. Presidential Executive Order on Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure, <https://www.whitehouse.gov/presidential-actions/presidential-executive-order-strengthening-cybersecurity-federal-networks-critical-infrastructure/> [↑](#footnote-ref-35)
36. https://www.supplychainbrain.com/blogs/1-think-tank/post/16330-why-cybersecurity-is-a-supply-chain-problem [↑](#footnote-ref-36)
37. **https://www.supplychainbrain.com/blogs/1-think-tank/post/16330-why-cybersecurity-is-a-supply-chain-problem** [↑](#footnote-ref-37)
38. **CERT-UK PAPER**  [↑](#footnote-ref-38)
39. Source: https://www.pivotpointsecurity.com/blog/steps-towards-vendor-risk-management-policy/ [↑](#footnote-ref-39)
40. NIST SP 800 [↑](#footnote-ref-40)
41. **CERT-UK White Paper,** [**https://www.ncsc.gov.uk/content/files/protected\_files/guidance\_files/Cyber-security-risks-in-the-supply-chain.pdf**](https://www.ncsc.gov.uk/content/files/protected_files/guidance_files/Cyber-security-risks-in-the-supply-chain.pdf) [↑](#footnote-ref-41)
42. NIST Special Publication 800-161, Supply Chain Risk Management Practices for Federal Information Systems and Organizations, Boyens et al, April 2015, https://doi.org/10.6028/NIST.SP.800-161 [↑](#footnote-ref-42)
43. NIST, Framework for Improving Critical Infrastructure Cybersecurity Version 1.1, [↑](#footnote-ref-43)
44. **https://csrc.nist.gov/CSRC/media/Projects/Supply-Chain-Risk-Managements/documents/nist\_ict-scrm\_fact-sheet.pdf** [↑](#footnote-ref-44)
45. ISO SP 800 [↑](#footnote-ref-45)
46. Source: https://www.pivotpointsecurity.com/blog/steps-towards-vendor-risk-management-policy/ [↑](#footnote-ref-46)
47. CERT-UK, Cyber-security risks in the supply chain [↑](#footnote-ref-47)
48. [**https://www.itgovernance.co.uk/iso27000-family**](https://www.itgovernance.co.uk/iso27000-family) [↑](#footnote-ref-48)
49. **https://www.sgs.com/en/health-safety/quality-health-safety-and-environment/risk-assessment-and-management/security-management/iso-28000-supply-chain-security-management-systems** [↑](#footnote-ref-49)
50. **http://www.eccinternational.com/** [↑](#footnote-ref-50)
51. CTIA: Wireless Industry Announces New Cybersecurity Certification Program, https://ctia.it/2wbVJzE [↑](#footnote-ref-51)
52. **ETSI report:**[**https://www.etsi.org/deliver/etsi\_gs/NFV-EVE/001\_099/005/01.01.01\_60/gs\_NFV-EVE005v010101p.pdf**](https://urldefense.proofpoint.com/v2/url?u=https-3A__www.etsi.org_deliver_etsi-5Fgs_NFV-2DEVE_001-5F099_005_01.01.01-5F60_gs-5FNFV-2DEVE005v010101p.pdf&d=DwMGaQ&c=RoP1YumCXCgaWHvlZYR8PZh8Bv7qIrMUB65eapI_JnE&r=fKXdw1OLG9OU7ouhOjOSuU0BYbM5xfWexnXIBHqx46k&m=VmTjU44xzorJUS1hiDEmUlVeelJHtiM2sznISs22h7I&s=5F7FGiSUzb5MeEvVHGz0a075SAOXUKGfMaZvtYwp6uU&e=) [↑](#footnote-ref-52)