SECURING 5G *and* EVOLVING ARCHITECTURES

PRAMOD NAIR

FREE SAMPLE CHAPTER

SHARE WITH OTHERS

Securing 5G and Evolving Architectures

Pramod Nair

Addison-Wesley

Boston • Columbus • New York • San Francisco • Amsterdam • Cape Town • Dubai • London • Madrid • Milan • Munich • Paris • Montreal • Toronto • Delhi • Mexico City • São Paulo • Sydney • Hong Kong • Seoul • Singapore • Taipei • Tokyo

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

The author and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact intlcs@pearson.com.

Visit us on the Web: informit.com/aw

Library of Congress Control Number: 2021917555

Copyright © 2022 Pearson Education, Inc.

All rights reserved. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, request forms, and the appropriate contacts within the Pearson Education Global Rights & Permissions Department, please visit www.pearson.com/permissions.

No patent liability is assumed with respect to the use of the information contained herein. Although every precaution has been taken in the preparation of this book, the publisher and author assume no responsibility for errors or omissions. Nor is any liability assumed for damages resulting from the use of the information contained herein.

ISBN-13: 978-0-13-745793-9
ISBN-10: 0-13-745793-6

ScoutAutomatedPrintCode

Vice President, Editorial
Mark Taub

Director, ITP Product Management
Brett Bartow

Executive Editor
Nancy Davis

Development Editor
Christopher A. Cleveland

Managing Editor
Sandra Schroeder

Project Editor
Mandie Frank

Copy Editor
Bart Reed

Indexer
Erika Millen

Proofreader
Donna Mulder

Technical Reviewers
Dave Hucaby
Keith O'Brien

Editorial Assistant
Cindy Teeters

Designer
Chuti Prasertsith

Compositor
codeMantra

Graphics
codeMantra

Pearson's Commitment to Diversity, Equity, and Inclusion

Pearson is dedicated to creating bias-free content that reflects the diversity of all learners. We embrace the many dimensions of diversity, including but not limited to race, ethnicity, gender, socioeconomic status, ability, age, sexual orientation, and religious or political beliefs.

Education is a powerful force for equity and change in our world. It has the potential to deliver opportunities that improve lives and enable economic mobility. As we work with authors to create content for every product and service, we acknowledge our responsibility to demonstrate inclusivity and incorporate diverse scholarship so that everyone can achieve their potential through learning. As the world's leading learning company, we have a duty to help drive change and live up to our purpose to help more people create a better life for themselves and to create a better world.

Our ambition is to purposefully contribute to a world where:

- Everyone has an equitable and lifelong opportunity to succeed through learning.
- Our educational products and services are inclusive and represent the rich diversity of learners.
- Our educational content accurately reflects the histories and experiences of the learners we serve.
- Our educational content prompts deeper discussions with learners and motivates them to expand their own learning (and worldview).

While we work hard to present unbiased content, we want to hear from you about any concerns or needs with this Pearson product so that we can investigate and address them.

- Please contact us with concerns about any potential bias at https://www.pearson.com/report-bias.html.

Credits

FIGURE	CREDIT/ATTRIBUTION
Figure 4-24	Courtesy of O-RAN Alliance e.V.
Figure 5-51a	Courtesy of Google Cloud
Figure 5-51b	Courtesy of Amazon Web Services, Inc.
Figure 5-51c	Courtesy of Microsoft Corporation
Figure 3-1	Anand R. Prasad, Alf Zugenmaier, Adrian Escott and Mirko Cano Soveri, 3GPP 5G Security, 3GPP, August 6, 2018
Figure 3-2	Anand R. Prasad, Alf Zugenmaier, Adrian Escott and Mirko Cano Soveri, 3GPP 5G Security, 3GPP, August 6, 2018
Figure 3-13	D. Hardt, Ed., The OAuth 2.0 Authorization Framework, Internet Engineering Task Force, 2012
Figure 10-3	Courtesy of Cisco Systems
Figure 10-4	Courtesy of Cisco Systems
Figure 10-11	Courtesy of Cisco Systems
Cover	Yurchanka Siarhei/Shutterstock

Dedication

I would like to dedicate this book to my family both near and far.
Thank you all for your unwavering support, motivation and patience
throughout the development of this book.

Table of Contents

Part III End-to-End 5G Security Architecture and Prioritizing Security Investments

Part IV Emerging Discussions

Foreword

Society is about to embark on a digital upgrade—the next generation of the world's mobile communication infrastructure—5G. Along with new and innovative capabilities, 5G also introduces new security features, vulnerabilities, and risks. 5G does not just represent significantly increased bandwidth and lower latency, but it is expected to fundamentally change the mobile ecosystem with new partnership models, network slicing, massive deployment of Internet of Things (IoT) devices, and ultimately, an increasingly critical dependency on the technology for society to function. Due to this, our ability to secure 5G will directly affect the resilience of critical infrastructure and national security.

Some of the security key risks affecting 5G confidentiality, integrity, and availability are supply chain risks, increasing complexity leading to new vulnerabilities, and inherent weaknesses in the standards. The supply chain risks have reached the geopolitical center stage due to the high societal impact of 5G, and this has led to national and EU-level regulations, risks assessments, and GSMA's accreditation scheme Network Equipment Security Assurance Scheme (NESAS). The inherent increased complexity of 5G leads to a wide range of new potential vulnerabilities that will require increased vigilance from product vendors, service providers, and users alike.

In order to manage these risks, 5G is equipped with a broad range of security features and capabilities, and GSMA has outlined a list of critically sensitive functions—virtualization infrastructure, controller, orchestrators, Internet gateways, network slicing, mobile edge computing, routing and switching of IP traffic at the core, database functions, authentication, and access control. As always, a security by design approach following a zero-trust approach, with secure deployments and good operational hygiene, is key to securing the world's 5G deployments.

In this book, Pramod Nair guides us through the evolution of cellular technologies from a security perspective, the security architecture, deployment modes and use cases of 5G, as well as discusses end-to-end security architecture and prioritizing security investments. His unique outlook as the Lead Security Architect, head of 5G security architecture in Cisco Systems, and from more than 20 years in security allows him to combine a theoretical and applied perspective for the benefit of both business and technical readers.

André Årnes, PhD
Senior Vice President and Chief Security Officer at Telenor Group
Professor II at the Norwegian University of Science and Technology

Preface

5G technology will redefine the way we perceive cellular networks and will touch almost every aspect of our lives. 5G is not about just being faster, bigger, or better; it's about enabling multiple services that we'll all consume on an everyday basis. It will give rise to a new ecosystem of developers building applications that exploit the openness of 5G to help you develop new use cases for consumption by enterprises and subscribers alike. New features in 3GPP Releases 16 and 17 help further enable new use cases for non-public deployment of 5G by industry verticals and tighter convergence of 3GPP and non-3GPP technologies, bringing in multiple deployment methods—including on-premises, hybrid, and fully public cloud-based deployments. The 5G ecosystem will see a breakout from 3GPP-only based architecture to an open, multi-technology, multi-standard, polyglot ecosystem.

This evolution of the technology landscape also requires an evolution of the security mindset. We should start thinking of security as a foundational layer. It should be one of the primary foundations for any planned 5G use case implementation. This requires embracing multilayered security beyond the requirements in 3GPP specifications.

The business operational risk, legal risk, and reputational risk exist not only for the companies providing 5G software and hardware infrastructure, but for all companies, nation-states, and individuals who provide and consume 5G technology.

The time is now to evaluate the cyber risk posture and apply innovative thoughts to how we can approach these challenges today and build for what's to come tomorrow.

Motivation for Writing This Book

Security in evolving cellular technologies is not an easy concept to grasp, as the technologies have evolved rapidly and are becoming increasingly complex and nuanced as they become more open, especially when you add 5G to the mix.

5G will also enable enterprises and industry verticals to deploy private 5G/non-public 5G networks (5G NPN networks) on their own, without any integration with service providers. This necessitates private and government sectors to fully understand the 5G threat surfaces, develop methods to mitigate the threats, and prioritize the investments in security.

The existing material on security and cellular technologies is dispersed across many resources and does not cover the end-to-end 5G threat surfaces, threat mitigation, examples of real-life deployment scenarios, and prioritization of security controls based on use cases and deployment scenarios. The learning curve for a person trying to understand the evolution in cellular technologies, new architectures, multiple deployment methods, threat surfaces, and mitigation techniques is extremely steep and sometimes unnerving.

It is not surprising that the topic of securing cellular technologies tends to flummox newcomers and even seasoned network security engineers.

This book brings all the information together and arranges the key topics in such a way that they can be easily consumed and understood. The main purpose of this book is to enable any person to understand the key aspects of securing 5G and evolving technologies. This book covers a range of topics; it will take you through the evolution of technologies from 2G to 5G, with deep dives into specific topics, such as securing non-public 5G networks/private 5G deployments and prioritizing security investments.

The goal of this book is to provide pragmatic views on securing 5G and evolving networks. The knowledge and information gathered through numerous customer workshops, brainstorming sessions with service providers, industry verticals, industry experts from multiple vendors, proof of concepts, and lessons learned from actual security deployments for 5G networks are detailed in this book. Discussions with multiple CSOs and CTOs have enlightened me on the key data points required for prioritizing security, which you will see highlighted in this book. Apart from service providers, industry verticals are expected to adopt 5G technology, and this area has been expanded into specific use cases, threats, and mitigation techniques. This book closes with a chapter discussing the key areas of security evolution that will motivate you to investigate different aspects of security as the network evolves. It is aimed at helping you create a new mindset while securing your networks of the future.

Who Should Read This Book

I have designed this book so that you can begin without any prior knowledge about 5G or any preceding cellular technologies. This book is written to be suitable for multiple levels of technical expertise, including the following:

- Security experts looking to understand the history of cellular technology evolution to 5G, key 5G security enhancements, and security challenges
- Early-in-career telecom engineers, transport design engineers, and radio engineers looking to design and implement mobile networks
- Government departments looking at security impacts of 5G deployment for use cases such as smart city and looking at implementing security measures
- Management consultants advising governments and service providers on 5G security strategy
- CSO and CTO teams from service providers looking at securing 5G deployments
- CSO and CTO teams from enterprises deploying NPN/private 5G
- Enterprise network design and implementation teams deploying NPN/private 5G deployments
- Security architects responsible for securing the mobile infrastructure
- Enterprise solution architects and enterprise security architects working with enterprises integrated with service provider 5G networks
- Security strategy teams within service providers, enterprise and industry verticals deploying 5G

- Cloud computing and data center teams involved with 5G strategy and deployment
- Enterprise solution and security architects deploying standalone private/NPN 5G or utilizing service providers' 5G slice network
- Audiences of varying levels of expertise from the military and defense community
- Audiences from industry verticals such as smart manufacturing, critical infrastructure entities and vendors, and autonomous vehicle manufacturers
- Cybersecurity vendor product managers looking for use cases or features to enhance security products to cater for secure 5G deployments
- Students who would like to get a quick understanding of cellular technologies and a look at the new features in 5G

Throughout the book, you will see practical examples and real-life scenarios of how you might architect a solution to mitigate threats and improve the security posture of your network.

How This Book Is Organized

To allow technical and nontechnical audiences to consume the book in an effective and organized way, it is split into four parts. The parts and chapters cover specific topics.

Part I, "Evolution of Cellular Technologies to 5G, Security Enhancements, and Challenges," explains the evolution of cellular technologies toward 5G as well as new security enhancements and new security challenges brought in by 5G. It will also take the reader through different deployment modes, including private 5G / non-public networks (NPN). This part will mostly cater to the audience who wants a high-level view of 5G technology and its security aspects. It includes the following chapters:

- **Chapter 1**, "Evolution from 4G to 5G," covers the evolution of cellular technologies and will provide you with a basic understanding of the 5G technology features. It will also take you through some of the key enhancements in 3GPP Rel-16 and Rel-17.
- **Chapter 2**, "Deployment Modes in 5G," covers the different non-standalone and standalone deployment modes and use cases, which can be mapped to specific deployment modes.
- **Chapter 3,** "Securing 5G Infrastructure," covers new security enhancements and new security challenges brought in by 5G. It also discusses the reasons why you should have an external layer of security controls, even though 3GPP provides some enhancements in security.

Part II, "Securing 5G Architectures, Deployment Modes, and Use Cases," covers the security controls for 5G network components such as RAN, transport, 5GC, and devices. It then takes you through securing 5G enablers—such as multi-access edge compute (MEC), software-defined networks (SDNs), network slicing, orchestration, and automation—and protecting different deployment methods such as on-premises, private and public cloud based MEC, and hybrid cloud, including open RAN deployments. It finally covers securing key 5G use cases such as critical infrastructure, vehicle-to-everything (V2X), and smart factory. This part of the book will be of keen interest to readers who would like to deep-dive into the security aspects of 5G and its key use cases. It includes the following chapters:

- **Chapter 4**, "Securing RAN and Transport Deployments in 5G," covers the 5G RAN and transport threat surfaces and threat mitigation for the 5G public and non-public deployments, including open RAN. This chapter also takes you through some real-world attacks and mechanisms to mitigate them.
- **Chapter 5**, "Securing MEC Deployments in 5G," covers various MEC deployment models, network functions deployed in the private and public cloud based MEC, its threat surfaces, and methods to mitigate the threats. The chapter also provides some real-world risk and risk mitigation scenarios.
- **Chapter 6**, "Securing Virtualized 5G Core Deployments," covers the threats due to virtualized 5G Core deployments and new methods of software development and deployment. This chapter also provides some key recommendations to secure your virtualized 5GC deployments with vendor-agnostic approaches and includes some real-world scenarios.
- **Chapter 7**, "Securing Network Slice, SDN, and Orchestration in 5G," covers network slicing and enablers of network slicing such as software-defined networks (SDNs), orchestration, and automation. The chapter also explains the threat surfaces and threat mitigations specific to network slicing and its enablers. This chapter also delves into the network slice as a service (NSaaS) offering, its threat surface, and methods to mitigate the threats.
- **Chapter 8**, "Securing Massive IoT Deployments in 5G," covers the risks related to IoT devices and related connectivity and management. The chapter then goes on to explain different security mechanisms and best practices to secure your network from any IoT device-based attacks.
- **Chapter 9**, "Securing 5G Use Cases," covers critical infrastructure, V2X, and smart manufacturing use cases, which use different types of IoT devices—some smart, some semi-smart—as well as non-smart devices. The chapter takes you through the risks within these three use cases and methods to mitigate the risks.

Part III, "End-to-End 5G Security Architecture and Prioritizing Security Investments," provides an overview of the various security recommendations for end-to-end 5G security and discusses the factors based on which certain security controls can be prioritized among other security controls for 5G networks. This part will be of keen interest to an audience who would like to have an end-to-end view of security and understand the methods to prioritize investments in security. It includes following chapters:

- **Chapter 10**, "Building Pragmatic End-to-End Security 5G Architecture," covers the key building blocks for creating an end-to-end security layer for 5G deployments. This chapter also provides you with a checklist for each of the 5G domains and includes zero-trust design principles.

- **Chapter 11**, "Prioritizing 5G Security Investments," covers the considerations and recommendations for prioritizing investments to secure your 5G network. This chapter takes two primary scenarios—one related to a service provider providing mobile service, and the other related to the non-public deployment methods for industry verticals and enterprises.

Part IV, "Emerging Discussions," takes you through the topics aimed at new features being discussed for 5G and evolving architectures, security enhancements using machine learning (ML) and artificial intelligence (AI), and the method to make your network quantum safe. This part will be of keen interest to readers who would like to understand the key discussions in the security industry around 5G and evolving technologies. It includes following chapter:

- **Chapter 12,** "5G and Beyond," covers the adoption and adaptation of 5G standalone technology with new use cases, convergence of non-3GPP and 3GPP technologies, application of AI and ML in securing 5G and evolving technologies, and the importance of deploying crypto-agile mobile networks.

Due to ongoing developments, Chapter 12 will occasionally be updated with relevant new content and insights on the book's website at www.informit.com. Register your copy of *Securing 5G and Evolving Architectures* on the InformIT site for convenient access to these updates and/or corrections as they become available. To start the registration process, go to informit.com/register and log in or create an account. Enter the product ISBN (9780137457939) and click Submit. Look on the Registered Products tab for an Access Bonus Content link next to this product, and follow that link to access any available bonus materials. If you would like to be notified of exclusive offers on new editions and updates, please check the box to receive email from us.

Please note that this book is written with a vendor-neutral approach, and it does not give recommendations on what vendor should be deployed. Each service provider or industry vertical planning to deploy 5G can evaluate the security controls required and make decisions based on their own criteria, circumstances, and targeted use cases. This book covers the details of the security controls, required features, and functions required for securing 5G and evolving networks, allowing you to make better informed decisions.

Happy reading, and I hope you enjoy reading this book as much as I enjoyed writing it!

Acknowledgments

I would like to acknowledge the tremendous support I received from the Cisco staff, especially my management team and colleagues.

Similarly, I would like to thank the reviewers, Keith O'Brien and David Hucaby, for their comments, feedback, and insights that enriched the content of this book.

I would like to thank Dr. André Årnes, PhD, Senior Vice President and Chief Security Officer at Telenor Group and Professor at the Norwegian University of Science and Technology, for writing the foreword.

I would like to extend my appreciation to the people from multiple companies who provided constructive comments during those numerous 5G security customer workshops and brainstorming calls.

I would like to thank executive editor, Nancy Davis, for her guidance, feedback, and massive support. I would also like to extend my thanks to the Pearson/Addison-Wesley team, especially to Chris Cleveland, the development editor, for his robust guidance throughout the editing process, and to Mandie Frank, for her support through the production process.

Finally, I would like to thank the many standards organizations, technologists, security experts, and industry peers who continue to contribute to the fields of both mobile communications and security.

About the Author

Pramod Nair is a Lead Security Architect at Cisco Systems focusing on service providers. During his 20 years of experience in the industry, Pramod has worked in multiple areas, including research and development, designing end-to-end mobile networks, and technical consulting on military and defense projects.

Among other responsibilities in his current role within Cisco, Pramod leads 5G Security Architecture, driving its adoption globally, and has been instrumental in architecting secure next-generation networks for customers across the globe. He is a regular speaker on the subject at large conferences and industry events.

Pramod is an active member of the security community. His role is to help mobile network providers, service providers, industry verticals, the national security and defense sectors, and other agencies dedicated to securing critical infrastructures. He is also deeply involved with industry trade organizations, has co-chaired the 5G security white paper within the 5GAmericas work group, and works with the National Institute of Standards and Technology (NIST) on 5G security.

Pramod holds a patent in fraud detection and has published various white papers and articles covering security-related topics.

Chapter 8

Securing Massive IoT Deployments in 5G

After reading this chapter, you should have a better understanding of the following topics:

- Threats in massive IoT use case deployments
- Securing massive IoT networks
- Real scenario case study examples of massive IoT threat surfaces and threat mitigation techniques

This chapter will take you through the threat surfaces in 5G massive IoT deployments and mechanisms to mitigate the threats.

This chapter will be of particular interest to the following teams from enterprise, industry verticals, Non-Public Networks (NPN), 5G service providers deploying 5G mIoT, and cybersecurity vendors planning product developments and new functionalities to secure 5G mIoT use cases.

- Mobile infrastructure strategy teams of service provider deploying mIoT in 5G
- Security strategy teams within service provider and enterprise verticals planning on deploying 5G mIoT
- Transmission and the packet core team within service providers and private 5G enterprises planning to deploy 5G mIoT
- Cloud computing and data center teams involved with 5G strategy and deployment
- Security architects and design teams looking at securing the public and non-public mobile infrastructure
- Solution and security architects deploying 5G mIoT on enterprises and industry verticals

- Enterprise solution and security architects using IoT services from mIoT service provider
- Government departments deploying 5G mIoT
- Cybersecurity vendor teams looking to secure mIoT deployments for their customers
- Product managers of cybersecurity vendors trying to identify use cases for new products or features to protect 5G mIoT deployments

5G represents a disruptive shift from just traditional consumer smartphones to advanced enterprise services, including ultra-reliable low-latency communication (URLLC)–based machine-to-machine (M2M) use cases. 5G is expected to be widely adopted in enterprise, industrial, and IoT use cases, enabling greater workforce mobility, automation, and countless new applications. Incorporation of 5G into these environments requires a deeper level of integration between end-user networks and 5G service interfaces, exposing both enterprise owners (in particular, operators of critical information infrastructure) and 5G service providers to new risks. Before we get into the risks and mitigation of risks, we will first need to look into the types of IoT use cases.

5G also sees a departure from the reliance on a single approach to authenticating all users onto the network-based SIM cards. The Third-Generation Partnership Project (3GPP) has addressed such shortcomings, with 5G now integrating the Extensible Authentication Protocol (EAP) framework, first adopted by Wi-Fi into WPA-Enterprise back in 2002, into its architecture. The 5G standard now provides examples of how to use EAP-TLS certificate-based authentication in 5G as well as other EAP methods that support mutual authentication. The list that follows outlines some of the key reasons why IoT threats are quite critical in 5G based on the excerpts taken from the Cisco Annual Internet Report (2018-2023):

- The number of devices connected to IP networks will be more than three times the global population by 2023. There will be 3.6 networked devices per capita by 2023, up from 2.4 networked devices per capita in 2018. There will be 29.3 billion networked devices by 2023, up from 18.4 billion in 2018.
- Globally, devices and connections are growing faster (10 percent compound annual growth rate [CAGR]) than both the population (1.0 percent CAGR) and the Internet users (6 percent CAGR). This trend is accelerating the increase in the average number of devices and connections per household and per capita. Each year, various new devices in different form factors with increased capabilities and intelligence are introduced and adopted in the market. A growing number of M2M applications, such as smart meters, video surveillance, healthcare monitoring, transportation, and package or asset tracking, are significant contributors to the growth of devices and connections. By 2023, M2M connections will constitute 50 percent of the total devices and connections.
- M2M connections will be the fastest-growing device and connections category, growing nearly 2.4-fold during the forecast period (19 percent CAGR) to 14.7 billion connections by 2023.

With this type of growth in the number of devices and spurts in new use cases such as M2M, an attack that successfully disrupts the network, or that steals or undermines the integrity of confidential data, could have a far greater economic and societal impact than previous generations.

IoT devices and applications have been around for quite some time and are not a new concept for 5G. There are networks today using LTE or NB-IoT technologies enabling IoT use cases. 5G offers flexibility in IoT deployment. The use cases aimed at 5G IoT are devices having different bandwidth requirements. Some require high bandwidth and transmit in burst, while some require low bandwidth and continuous connectivity. 5G offers this capability to support the massive number of devices with different bandwidth requirements. In addition, 5G also supports enterprise and industry use cases that have strict requirements on latency. This is one of the key reasons why the industry is looking at adopting 5G. The flexible mode of 5G deployment using network slicing and deployment of applications in the edge of the network can bring down the latency to 1ms or less, enabling ultra-reliable and low-latency use cases such as factory automation, enhanced vehicular technologies such as vehicle-to-everything (V2X), power and utility sector use cases such as smart energy grids, and other demanding use cases to become a reality.

There are different types of IoT use cases in 5G depending on the data consumption, energy consumption, and scale of deployment. When you take a step back and look at the use-case scenarios in 5G, we can split the IoT devices into smart devices and not-so-smart devices. Smart IoT devices are the devices that have some intelligence built into them and can make some decisions based on the input data. The not-so-smart IoT devices are the devices that just send the collected data and receive certain actions, such as stop data collection and a query to start data collection.

Use cases attributed to 5G such as smart cities would require the use of both types of devices, as shown in Figure 8-1, and have an artificial intelligence (AI), machine learning (ML), and an analytics layer to analyze the information from multiple devices and make a decision based on it. An example could be automated car parking in a busy area such as an airport parking lot, as shown in Figure 8-1.

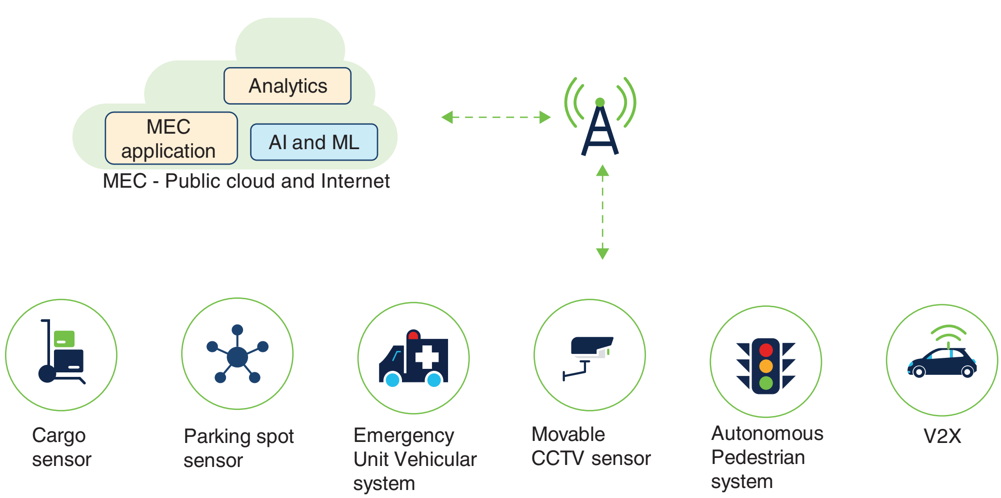

FIGURE 8-1 Different IoT Device Types to Enable a 5G Smart City

As shown in Figure 8-1, it would require different types of mIoT devices to enable the smart city use case. Table 8-1 lists the types of devices to fulfill the use case of finding a parking spot and the safest way to reach the parking spot.

TABLE 8-1 Different IoT Device Types

IoT Device	mIoT Device Type	Function
Cargo sensor	Not-so-smart device	Sends the geo-location metadata along with the speed
Parking spot sensor	Not-so-smart device	Indicates whether or not a vehicle is located in a parking spot
Emergency Unit Vehicular system	Not-so-smart device	Indicates whether an emergency vehicle is active in the location
Movable CCTV sensor	Not-so-smart device	Detects if there is movement near the parking spot
Autonomous pedestrian system	Smart device	Indicates any V2X application in the vicinity and broadcasts a message based on whether or not a pedestrian is crossing. Captures any speeding instances and sends data to the road safety officers. Indicates any collision and immediately broadcasts messages to the emergency health unit.
V2X	Smart device	Provides a road safety application such as intersection movement assist, provides emergency brakes, and also includes V2V (vehicle-to-vehicle) communications

As listed in Table 8-1, to fulfill this example of smart city–based parking, there is a need for both not-so-smart-devices and smart devices.

In this example, the cargo sensor, Emergency Unit Vehicular system, and autonomous pedestrian system are all part of the collision-prevention mechanism. The parking spot sensor and movable CCTV sensor are part of the parking detection mechanism. The V2X system is embedded within the vehicle for passing along the metadata to the MEC application.

All the data from the mIoT devices is then passed on to the AI and ML system and real-time (RT) analysis system. The AI, ML, and analytics system will then detect the free parking spot and the safest way to approach the parking spot and then help park the car or indicate the parking spot and the best way to reach it.

Massive IoT in 5G addresses the need to support billions of connections with a range of different services. IoT services range from device sensors requiring relatively low bandwidth to connected cars that require a similar service to a mobile handset. Network slicing provides a way for service providers to enable services to enterprises, giving them the flexibility to manage their own devices and services on the 5G network. mIoT, as the name suggests, is a category of use cases that is driven by scale.

Figure 8-2 illustrates an example of components that are part of the mIoT deployment.

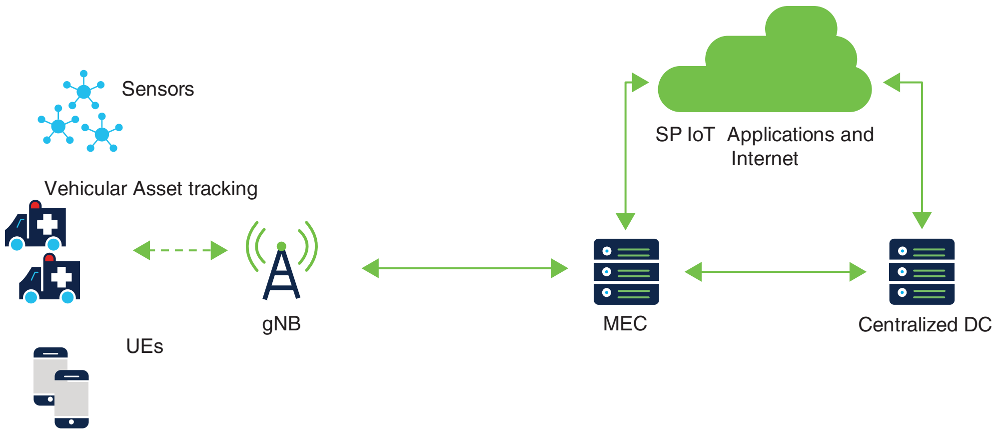

FIGURE 8-2 mIoT Deployment in 5G

Figure 8-2 shows an example of mIoT use-case deployment using 5G. The gNB serves geographically disparate devices such as sensors and vehicles that need to be tracked. mIoT would typically include devices that transmit and consume low data and are in the scale from hundreds to millions. Depending on the device type, it could be low-energy-consuming devices with limited access to power with a very light software stack for communications. There are device vendors in the market with 5G-capable chips with optimized power consumption.

This chapter will cover the 5G MIoT part. 5G IoT use cases based on smart devices (V2X, smart city, industrial IoT use cases, and so on) are covered in Chapter 9, “Securing 5G Use Cases.”)

Massive IoT–Based Threats in 5G

Figure 8-3 shows the key threats for the device-based threats for the devices connecting to the service provider’s 5G infrastructure. The devices in this case can be the 5G user equipment (UE), sensors, and IoT devices connecting to the 5G network provided by the service provider.

Figure 8-3 shows 5G multi-access edge compute (MEC), centralized 5GC (5G Core), public or private cloud-based SP applications, and the Internet access layer. Depending on the deployment plans of the service provider, the 5G User Plane Function (UPF) would be deployed in the MEC, along with any of the IoT applications that require caching. When the UPF network functions are deployed in the MEC, the N6 interface—the interface between the data network (DN) and the UPF—is also configured to allow UE and 5G devices to interconnect with the data network. Depending on the deployment scenario, the 5GC could host the 5G network functions that have low impact with higher latency, such as control plane functions, user plane functions for some IoT use cases, and the operations, administration, and maintenance (OAM) functions. Many service providers are also planning to have the configuration management (CM), fault management (FM), and performance management (PM) for the consumer IoT devices being catered to from the public/private cloud.

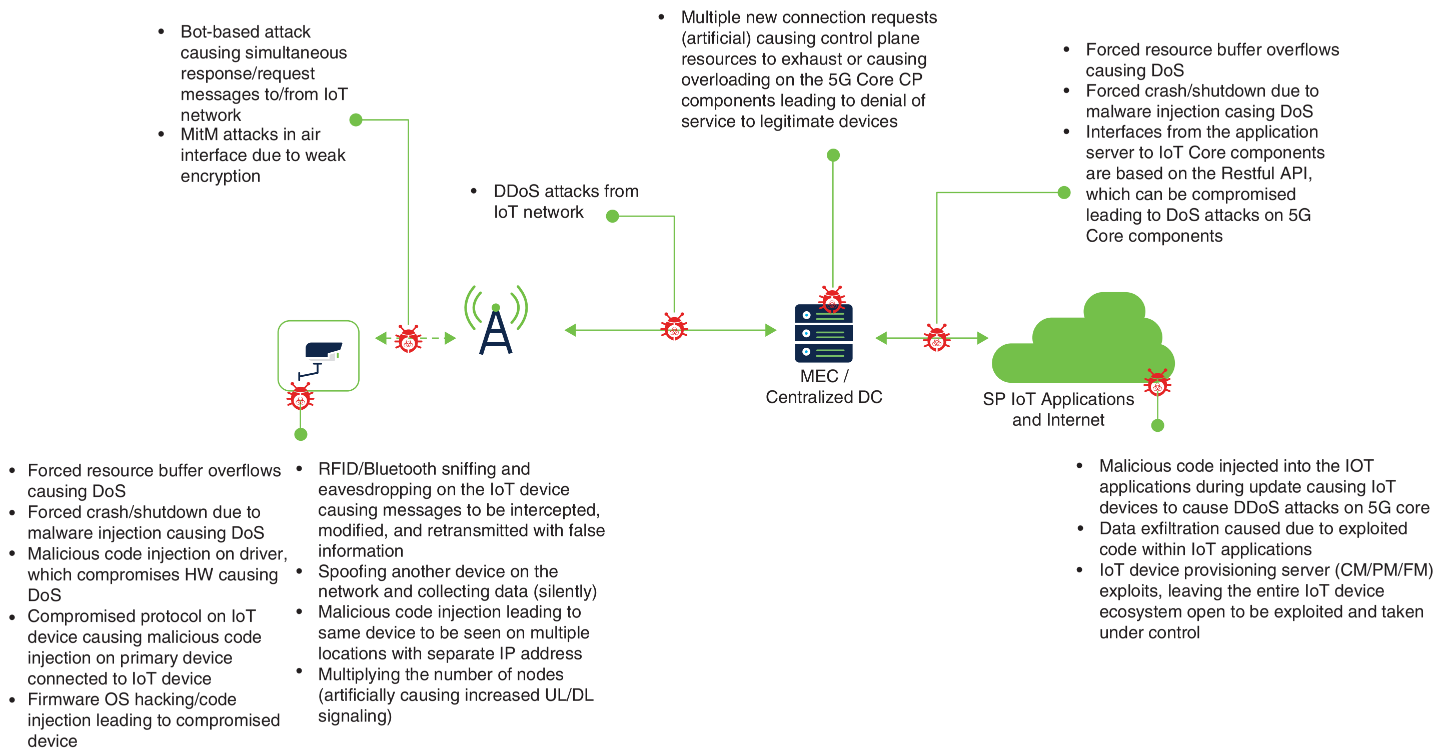

FIGURE 8-3 mIoT Threat Surface in 5G Deployments

The majority of the threat surfaces illustrated in Figure 8-3 are primarily due to the device vulnerabilities and the devices being compromised by the command and control (C&C) server.

Here are some of the key threats related to mIoT use cases within the 5G networks:

- C&C-based attacks
- Malicious code injection on the driver that compromises the hardware, causing a denial of service (DoS)
- Forced resource buffer overflows causing DoS
- Forced crash/shutdown due to malware injection, causing DoS
- Compromised protocol on an IoT device, causing malicious code injection on the primary device connected to the IoT device
- Firmware OS hacking/code injection, leading to a compromised device
- Radio-frequency identification (RFID)/Bluetooth sniffing and eavesdropping on the IoT device, causing messages to be intercepted, modified, and retransmitted with false information
- Spoofing another device on the network and exfiltrating data
- Malicious code injection leading to the same device being seen at multiple locations with separate IP addresses
- Multiplying the number of nodes (artificially), causing increased signaling in both UL/DL

Device Vulnerabilities Due to Weak Built-in Security

mIoT devices usually have very weak built-in security mechanisms due to lower price points of the devices to make them affordable to a large consumer base. The IoT deployment of any type, be it based on smart IoT devices or not-so-smart IoT devices, needs to be catered to by robust security controls to mitigate the vulnerabilities introduced by weak built-in security mainly due to the low cost and limitations due to the form factor. Non-mIoT use cases that are not geographically located would also need multilayered security controls to secure them from targeted attacks very specific to industry verticals, such as major automotive manufacturers or government utility verticals.

Spoofing, cloning, and eavesdropping on the 5G endpoints/IoT devices can be carried out by attackers impersonating an RFID or Bluetooth device and reading and recording the transmitted data from the 5G-enabled IoT device. This is primarily made possible due to weak access controls and poor authentication methods used by the IoT device. These kinds of attacks are more prevalent in verticals of IoT such as healthcare where the IoT devices use Bluetooth to transfer the patient's health statistics to a tablet where the vital stats of the patient can be checked/monitored by the healthcare workers.

Another type of attack mentioned in Figure 8-3 is where the devices are compromised. In this instance, all the data from the impacted devices is dropped or redirected instead of being transmitted to the intended receiver for further forwarding or analysis. The data from such devices can then be analyzed by the attacker for any valuable data points, such as the IP address of the receiver, which can then be targeted for DoS.

These kinds of attack methods can also be referred to as *sinkhole attacks* or a form of *routing attack*. This is because the method of attack used in such instances is to route the packets away from the main intended receiver. To prevent the detection of such attacks, the data can be mirrored to the malicious data collection server using a method very similar to port mirroring or Switch Port Analyzer (SPAN), which is used quite commonly in the network monitoring environment of the service provider networks. SPAN copies (or mirrors) traffic received or sent (or both) on source ports or source VLANs to a dedicated destination switch port for analysis. You can analyze network traffic passing through switch ports or VLANs by using SPAN or Remote SPAN (RSPAN) to send a copy of the traffic to another port on the switch or on another switch that has been connected to a network analyzer or other monitoring solution.

Management layer–based attacks are another key concern for device-based attacks within 5G. In these attacks, the attacker tries to take control of the key management layers, such as CM, FM, and PM, by exploiting the existing vulnerabilities of the IoT vendors' management platform or the open source components used in the vendors' IoT platform. Once the vulnerability has been successfully exploited, the attacker gains access and control over all endpoints catered for by the IoT vendor for the service provider. This can now be used for DoS and distributed denial-of-service (DDoS) attacks. One of the methods the attacker could also use here is to change the encryption type or level (from encrypted to null encryption), which makes the entire IoT network susceptible to man-in-the-middle (MitM) attacks.

The key threat surfaces and vulnerabilities are discussed in more detail in the sections that follow.

Supply Chain Vulnerability

Supply chain vulnerability is a well-known issue across different industry segments. The challenge of supply chain vulnerabilities becomes more prominent in 5G, as it enables attaching millions of low-cost IoT devices to the network. 5G also introduces critical infrastructure–based use cases and caters for use cases like smart cities, defense, and so on. These critical infrastructure 5G IoT use cases attract more nation-state attackers and thus are under higher levels of risk for cyberattacks. Supply chain is one of the weak links in security. If not secured properly, it opens the door wide for attacks, and the impacts of the attacks could be devastating, depending on the use case where the vulnerable IoT device was used. This section will take you through the vulnerabilities in the IoT supply chain related to manufacturing and distribution, as shown in Figure 8-4.

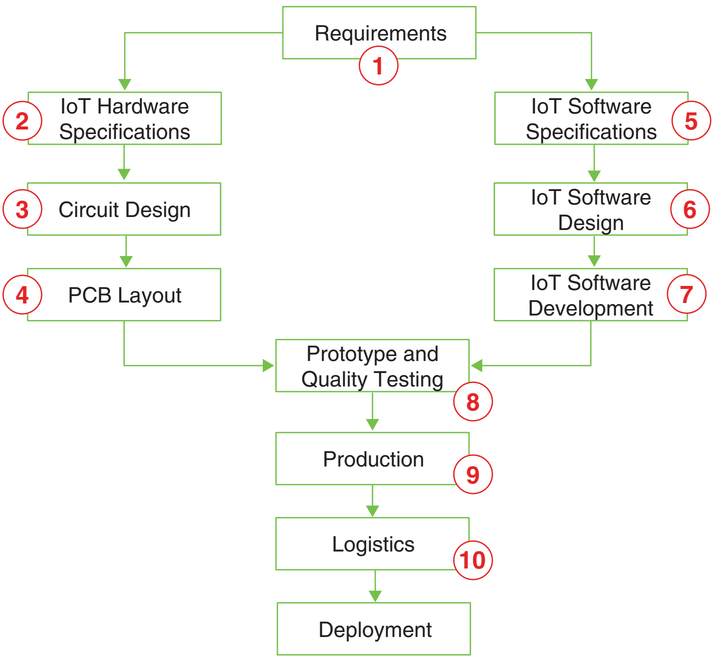

FIGURE 8-4 Vulnerabilities in Different Stages of the Supply Chain

Key vulnerabilities and threat vectors for the IoT supply chain related to manufacturing and distribution are explained in the list that follows:

1. The requirement stage is when you send the requirements for your IoT device to the vendor. This will include details like maximum energy consumption, dimension of the unit, maximum/minimum temperature, pressure (depending on use case), software or platform requirements such as integration options using API, and so on. The threat vector here is the requirement that is actually passed on to the vendor product R&D and manufacturing team. An attacker might add a couple of details in the requirements not actually requested by you. These newly added details are aimed at creating the backdoor using hardware or software remodifications to the original design, which can then be exploited by the attacking entity once deployed.

2. The hardware specification team would normally take the requirements from the customer and map them to the required hardware, including deciding what sort of components should be used in manufacturing the device. Typical considerations are values to withstand humidity, temperature, power consumption, and so on. The threat vector here is that an attacker could choose certain components that will fail when a certain condition is met. For example, the malicious actor or the attacking entity could intentionally choose a substandard electronic component or a customized component that fails after a certain temperature or humidity level is reached.

3. Once the components are finalized, the design team would make a schematic of the design that will be used as a blueprint for the printed circuit board (PCB) manufacturing for the IoT device. This is a very important part of the manufacturing process, as all the further checks on quality and so on would be referred back to the schematic. The attacking entity or the malicious actor could alter the design to include an eavesdropping component to leak sensitive data to a predetermined destination such as a C&C server.

4. The PCB layout process and component soldering are the next steps after the circuit design process. Here, the key vulnerabilities and threat vectors are due to the attacker choosing counterfeit electronic components causing intermittent failures that are difficult to find and correct.

5. IoT software specifications are taken from the requirements list you have provided to the IoT vendor/manufacturer. A member of the IoT software specification team or an attacker working in the software specification team could be directed to modify the specification for the software. The software specification will also be used in the software quality process for validating the software and to ensure that the designed software meets the software specifications. Any modification done in the software specification process will be considered as the software blueprint for the device.

6. The software design team would follow the specifications set by the software specifications team and specify the architecture and software technology to be used. In this process, the vulnerabilities are mainly due to the lack of knowledge about security leading to weak software for the device.

7. The software development team programs the IoT device with the chosen software language. With attacks aimed at software vulnerabilities on the rise, it is imperative that the software team follows secure software design and avoids known vulnerabilities such as buffer overflows, which occur when there is more data in the buffer than it can handle, leading to software crash and thus creating a point for cyberattack. This can be intentionally implemented by an attacker within the software development team. Another threat vector is when a team member of the software development team is instructed by an attacker or an attacking entity to include malicious code within the program to allow a backdoor entry to the device or to the private network where the IoT device is deployed.

8. In the post-PCB layout and software development process, the IoT device manufacturer would validate whether the hardware prototype and software fulfill the requirements set by your (or your customer's) IoT device requirement. This is the last part of the process when a vulnerability can be identified and patched. If the quality team is compromised by an attacker, the specific vulnerability that is planned to be exploited by the attacker/attacking entity will be overlooked and will not be patched. This will leave the IoT device open for any attacks.

9. One of the key vulnerabilities in production is shadow production. Shadow production is where the real production numbers are hidden and used to flood the market with IoT devices with backdoors and vulnerabilities, making the devices open to attacks. Another threat vector is where the Joint Test Access Group (JTAG) ports are left unsecured. JTAG is an interface

that provides an option for debugging, reprogramming, and so on. In many gaming consoles, the JTAG ports are unsecured and open to user access. If you had the common interface cable for JTAG, you could plug it into your computer, use manufacturer default credentials, and play pirated games with some modifications on the attributes using the JTAG ports. The same unsecured JTAG port in an IoT device can allow an attacker to have unauthorized access and possibly have access to the private network where the IoT devices are deployed. The physical attacks, such as injecting malicious code into the IoT network, can be made possible by tampering with an IoT endpoint, gaining control over it, and then using that endpoint to gain access into the central IoT network. Attackers also exploit the JTAG interface used by manufacturers for debugging purposes. JTAG is an industry standard for on-chip instrumentation in electronic design automation (EDA). JTAG is also used to program field-programmable gate arrays (FPGAs). Most CPU vendors still use JTAG for debugging purposes. If JTAG ports are left unprotected, this interface can become a critical attack vector on the system.

10. Logistics is the other vulnerability in the supply chain that is prone to sabotage or modification of the IoT devices while in transit. Though this is not the most preferred attack vector for IoT devices in the supply chain, for critical infrastructure use cases, logistics needs to be carefully monitored. Your supply chain risk management (SCRM) should ensure that you have the right controls, such as choosing validated and security-cleared logistics vendors for shipping and transportation of IoT devices from production to deployment.

The attacks are primarily aimed at data exfiltration, tampering with the files within the IoT network, and gathering information. With the control garnered over the IoT network, the attacker could control the operations and the data flow between the IoT network and the 5G network components, such as a radio (gNB) or storage/configuration in the MEC layer of the 5G network. With the control over the IoT network, the attackers can damage the IoT devices and disrupt the IoT service, thereby causing DoS to service providers' IoT services. This is not a new threat vector for 5G technology specifically; it is prevalent in legacy technologies such as 2G, 3G, and 4G, but it's critical for 5G technology, as it is aimed at enabling IoT use cases such as mIoT that would impact different government and private sectors.

Command-and-Control Servers and Botnets

A command-and-control server (also referred to as a C&C, C2C or C2 server) is an endpoint/device that is compromised and controlled by an attacker. Devices on your network can be commandeered by a cybercriminal to become a command center or a botnet (a combination of the words "robot" and "network") with the intention of obtaining full network control. Establishing C&C communications via a Trojan horse is an important step for attackers to move laterally inside service provider networks, infecting machines and servers with the intent to exfiltrate data.

One famous example of botnet malware is Mirai, which causes its infected devices to scan the Internet for the IP address of IoT devices by using a table of common factory-default usernames and passwords. The Mirai malware then logs in to the IoT devices and infects them with the Mirai malware.

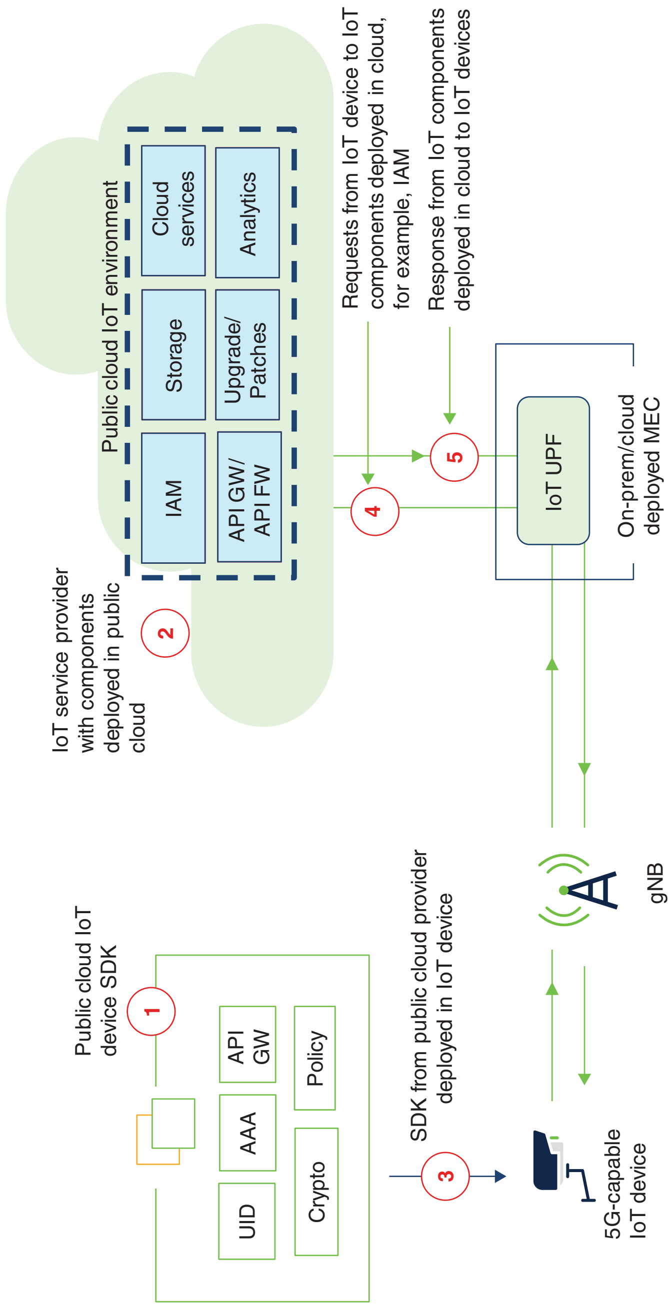

FIGURE 8-14 IoT Device Integration with Public Cloud Using SDK

Figure 8-14 illustrates the authentication of the IoT device using the installed SDK and is explained as follows:

1. The SDK will include open source libraries. The recommended practice for low-powered devices is to use an SDK that supports device connections that use Message Queuing Telemetry Transport (MQTT). The SDK will include a basic set of functionalities and policies to access the cloud-based IoT provider.
2. The key functionalities of the IoT service provider are deployed in the cloud. One of the key components is Identity and Access Management (IAM), which is used for authenticating the IoT devices. The API gateway (API GW) is used to protect the IoT applications from API vulnerabilities, such as providing rate-limiting functionalities and enhanced authentication and authorization functions.
3. Installing SDKs in the IoT devices will help you integrate IoT products to your choice of IoT providers deployed in public cloud.
4. The SDK deployed within the IoT device will initiate an HTTPS request toward the authentication, authorization, and accounting (AAA) component of the cloud-based IoT provider. The HTTPS request includes the X.509 certificate, which is verified by the AAA component to authenticate the IoT device.
5. Once the mutual authentication is performed, initial configuration can be downloaded to the IoT device. One of the other functions that can be performed is to attach a policy for the device, such as allowing the device to connect to the analytics engine, enabling you to enhance the services being offered to the IoT use cases.

In pragmatic deployment considerations, you also need to consider integration of hundreds of thousands or even millions of devices, which might require AAA to be deployed in the public cloud, as illustrated in Figure 8-15.

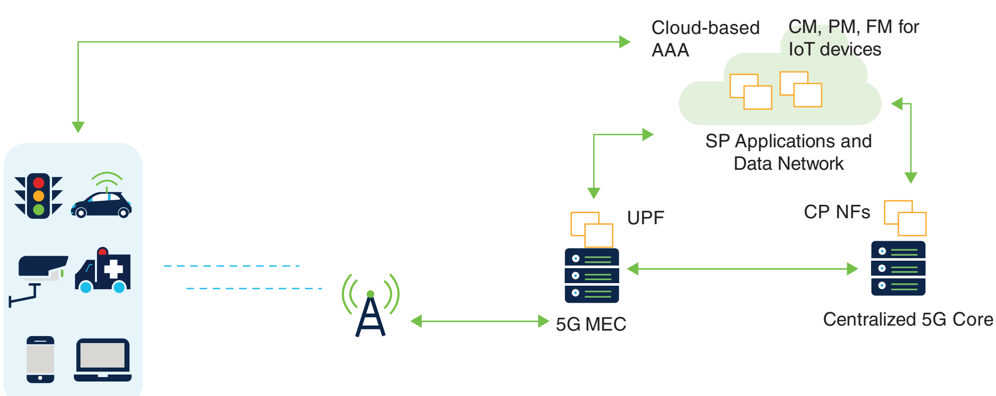

FIGURE 8-15 Cloud-Based Authentication for IoT Devices

One way to tackle the issue of identifying millions of devices is to build a strategy around having a unique ID (UID) assigned during the manufacturing process that can be used to identify and authenticate the device. Having a unique ID will also allow service providers to have proper lifecycle management, including tracking the software and hardware changes. Any infection or abnormal behavior can be easily tracked down to a specific device or group of devices.

Network Slice Isolation and Segmentation

Network slicing is one of the key evolutions of the network deployment brought in by 5G technology. Network slicing is the ability of the network to (automatically) configure and run multiple logical networks as virtually independent business operations on a common physical infrastructure. Network slicing is a fundamental architecture component of the 5G network, fulfilling the majority of the 5G use cases. Many operators are considering the offer of a network slice per enterprise, which is not that dissimilar to the per access point name (APN) offer for an enterprise in play today. As we consider the points where the enterprise then touches the 5G slice, a number of security aspects must be addressed—one of them being slice-level isolation, as illustrated in Figure 8-16.

Network slicing architecture, which allows the ability to run multiple logical networks as virtually independent business operations on a common physical infrastructure, also requires high isolation between the slices. Isolation within the components of the slice prevents the vulnerabilities from spreading to other components within the slice and between the slices in the case of any malicious attacks.

Intra-slice and inter-slice isolation should be implemented for both public and non-public networks (NPNs). The network slices should also allow a quarantine slice for identified malicious hosts, which provides isolation and restricts the spread of malware due to lateral movement.

Intra-slice can be provided by ensuring that the CNFs serving the slice are deployed on separate hosts. This ensures high availability for the slice.

Inter-slice isolation can be provided by deploying 5GC CNFs on separate hosts and then implementing network segmentation between slices. This mitigates malware propagation between slices of different sensitivity, such as a slice serving critical infrastructure (considered a highly sensitive slice) and a slice serving IoT devices (considered a less sensitive slice).

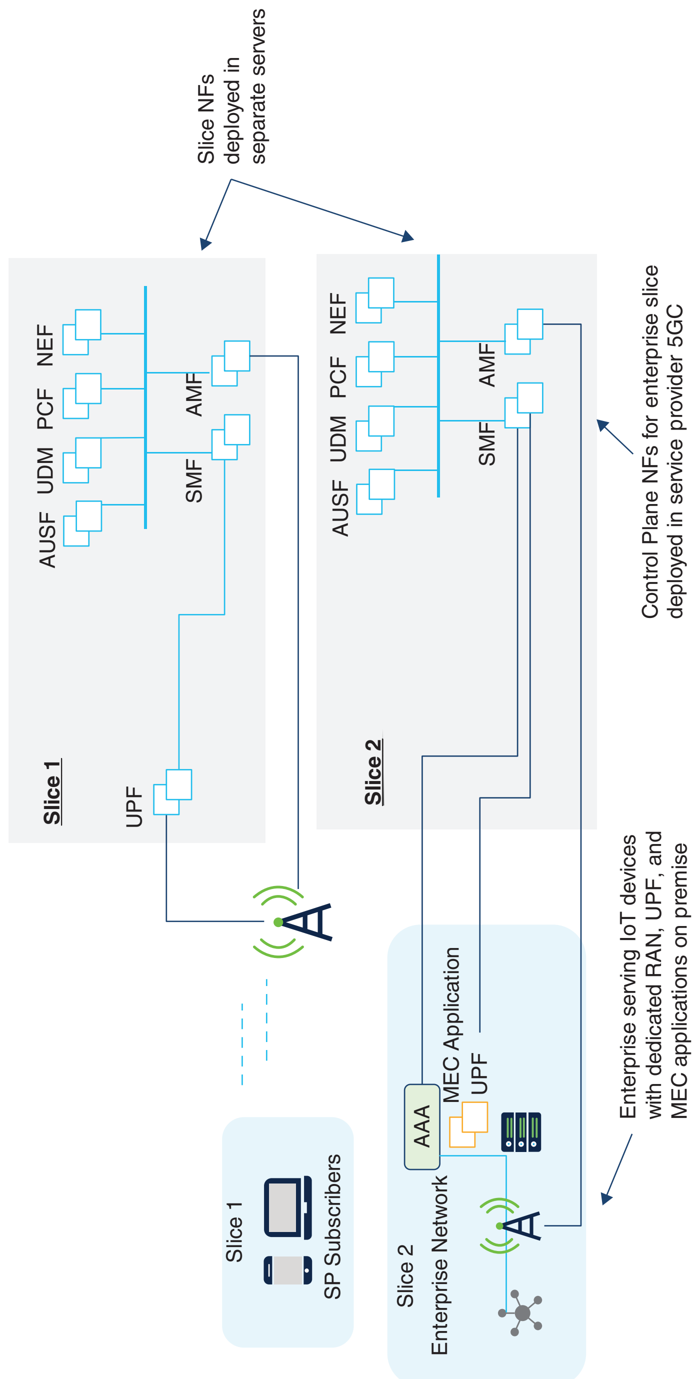

FIGURE 8-16 Slice Isolation and Segmentation for IoT Devices

Segmentation and isolation mechanisms used for the IoT deployment will vary depending on your deployment mode to cater for the mIoT use cases. If network slice mechanisms are used to provide access to the IoT device, you should ensure that the 3GPP 5G functions are isolated from other slices. This can be done by using separate x86 servers for deploying mIoT slice NFs. You should also architect your network such that web-facing applications are in a separate security zone and are not deployed in the same x86 server. This will ensure physical separation of the NFs and will reduce the probability of any side-channel attacks exploiting the vulnerability of the host OS and hardware (HW). If the mIoT devices are being deployed in the NPN network, then you should ensure that you have the mIoT network and the operational technology (OT) completely isolated from your IT network using a demilitarized zone (DMZ). In fact, if the mIoT is being deployed for critical use cases, there should be integrations with the IT network only if it is really necessary. Remote access to such networks should follow stringent identity and access mechanisms and should be continuously audited. This could be done by using a next-generation firewall (NGFW) integrated with your Network Access Control (NAC) and IAM layers.

Securing network slices is covered in detail in Chapter 7, "Securing Network Slice, SDN, and Orchestration in 5G."

Mitigating IRC and P2P-Related Attacks

In general, deploying IRC and P2P IoT devices in the subscriber's location should be avoided. But pragmatically speaking, it is well known that the security team of the service provider is rarely informed of IoT devices being sold to customers by the customer-facing teams of the service provider. To solve this issue, recommended practice dictates that service providers check the type of device, secure the development lifecycle followed by the device manufacturer, and look at the supply chain lifecycle of the device manufacturer.

If the existing devices within the service provider use IRC, then in cases of IRC-related botnet attacks, each bot client must know the IRC server, port, and channel. Anti-malware solutions available today can detect and shut down these servers and channels, effectively halting the botnet attack. If this happens, clients are still infected, but they typically lie dormant since they have no way of receiving instructions. A botnet can also consist of several servers or channels. If one of the servers or channels becomes disabled, the botnet simply switches to another. It is still possible to detect and disrupt additional botnet servers or channels by sniffing IRC traffic, which can be catered for by anti-malware and monitoring solutions.

If the existing devices within the service provider use P2P, for mitigating the P2P attacks that use the firewall pin-holing technique, then granular firewall configurations to block traffic on specific ports should be used. This would prevent infected devices from communicating with the malicious P2P servers.

Zero-Touch Security

Many of the consumer devices aimed at enabling IoT use cases use Zero Touch Provisioning (ZTP) to allow the PnP capabilities. This is done to allow easier deployment for the customer and provide a better user experience. Before choosing such devices from a manufacturer or vendor, the service provider should check whether the device manufacturer or vendor uses ZTS as a model for the ZTP process. Depending on the vendor, the method of ZTS is also called *secure zero touch* or *zero touch secure identity*, or other variants.

Implementing ZTS by the device vendor is quite critical, as it secures the device and authenticates and encrypts its communication with the cloud-hosted provisioning and configuration server or PnP servers and provides a secure lifecycle thereafter, including secure auto-deployment of patches, secure auto-installation of updates, and so on.

ZTS techniques used by the vendor should also ensure continuous authentication if any anomaly in behavior is detected or if reauthentication of the device occurs at certain intervals without interrupting the device functions. During assessment of the device vendor by the service provider, scalability of the solution should also be verified. Quite a few vendors in the market today use artificial intelligence (AI) to detect anomalies in behavior and can initiate the detection and response capabilities automatically depending on the behavior of the devices, including triggering the reauthentication of the devices and moving the devices with anomalous behavior to an isolated segment.

DNS Security for 5G IoT Devices

The Domain Name System (DNS) plays a very important role in the IoT ecosystem. The 5G devices enabling consumer IoT would primarily be using cloud-based provisioning servers for PnP, which is usually configured using an FQDN that will have the URL of the provisioning server configured or hard-coded. Using this configuration, the device will connect to the provisioning server, get authenticated (depending on the device vendor), and then connect to cloud services to transmit and receive data.

One of the key threats is DNS cache poisoning attacks, where a malicious or fraudulent IP address is logged in the local memory cache. The device configuration can also be modified for it to connect to a malicious server. This is because the devices trust the domain names to be secure. If an attacker changes the original domain name within the configuration template of the device or can change the hardcoded domain name to a malicious one, the device will try connecting to that domain name. The attacker can then insert a rogue update to the device, potentially taking full control of the device and targeting it against the service provider infrastructure, causing a DDoS attack or taking down the infrastructure, causing a DoS attack.

DNS, although scalable, does not include any inherent security mechanisms such as encryption, which makes it vulnerable to MitM attacks for interception and manipulation. Domain Name System Security Extensions (DNSSEC) and DNS over HTTPS (DoH) improve the security capability of DNS. DNSSEC is becoming more important for IoT devices due to the fact that it secures parts of the supply chain system as well. When an IoT device is manufactured, many of the device vendors use the cloud-based configuration for shipping and initial factory configuration. This is because many of the orders from service providers can be customed labeled so that when the customers receive their devices, they will be in the name of the service provider. This requires some changes at the manufacturing end, and many of these processes are automated in the industry these days. Secure DNS solutions can also be used by the service providers to enhance security for the IoT devices. This is further explained in detail in this section.

DNSSEC

DNSSEC is a set of extensions to DNS that provides a security chain of trust and protection from DNS vulnerabilities. DNSSEC provides DNS clients with cryptographic authentication of DNS data by using cryptographic keys to validate connections between the DNS client and a domain name.

Having DNSSEC as part of the device capability will ensure that the device is routed and connected to the authentic server.

Although DNSSEC adds integrity and trust to DNS, it does not provide confidentiality (DNSSEC responses are authenticated but not encrypted), which means that the DNSSEC responses can be intercepted. As the attacker can attempt to use DNSSEC mechanisms to consume a victim's resources, it does not provide complete mitigation against DoS attacks.

DoH

DNS over HTTPS (DoH) caters for DNS resolution using the HTTPS protocol. Using HTTPS, DoH provides better user privacy and prevents MitM-type attacks because it includes encryption between the DoH client and the DoH-based DNS resolver. DoH is published by the IETF as RFC 8484.

DoH works just like a normal DNS request, except that it uses Transmission Control Protocol (TCP) to transmit and receive queries. DoH takes the DNS query and sends it to a DoH-compatible DNS server (resolver) via an encrypted HTTPS connection on port 443, thereby preventing third-party observers from sniffing traffic and understanding what DNS queries users have run or what websites users are intending to access. Because the DoH (DNS) request is encrypted, it's even invisible to cybersecurity software that relies on passive DNS monitoring to block requests to known malicious domains.

If service providers plan to use DoH-based endpoints, there are certain mechanisms the security team can put into place to ensure that the devices use specific browsers. Browsers such as Chrome ensure that DoH will only be enabled when system DNS is observed to be a participating DNS provider. After DoH is enabled in Chrome, the browser will send DNS queries to the same DNS servers as before. If the target DNS server has a DoH-capable interface, then Chrome will encrypt DNS traffic and send it to the same DNS server's DoH interface.

Secure DNS

In many cases, consumer IoT devices today are not yet fully DNSSEC or DoH capable. One of the mitigation mechanisms from DNS cache poisonings and malicious DNS configurations is to use a cloud-based DNS security layer that ensures that the DNS request is not resolved to a malicious domain. There are vendors in the market today that integrate the secure DNS resolution along with the threat intelligence, anti-malware, and antivirus capabilities.

As illustrated in Figure 8-17, when the DNS security layer receives a DNS request from a 5G-capable IoT device, be it for the provisioning or PnP layer or for CM, PM or FM, it should use threat intelligence to determine if the request is safe, malicious, or risky—meaning the domain contains both malicious and legitimate content. Safe and malicious requests can be routed as usual or blocked, respectively. Risky requests can be forwarded to an inspection layer for deeper inspection. The secure DNS layer should also inspect the files attempted to be downloaded from the sites using antivirus (AV) engines and anti-malware protection, and based on the outcome of this inspection, the connection should be either allowed or blocked.

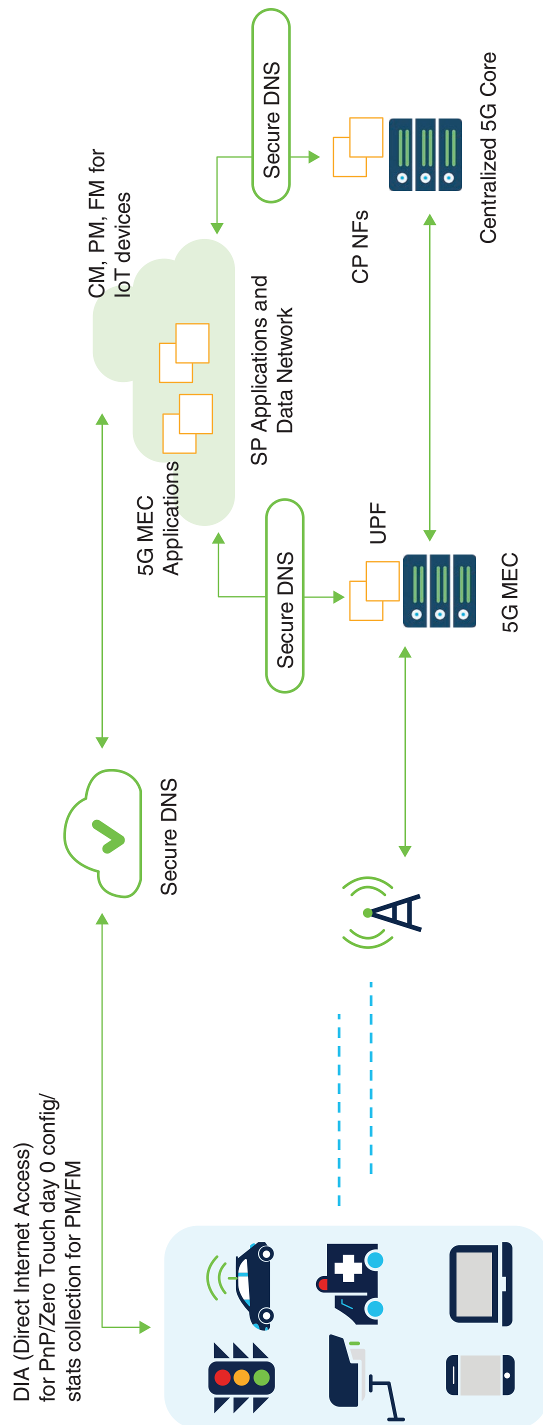

FIGURE 8-17 Securing DNS Layer Communication of IoT Devices

This is one of the most effective methods that will lead the security teams to remediate fewer instances of malware, and the threat is mitigated even before the devices are impacted or an attack is launched. Service providers selecting vendors or partners for secure DNS solutions should ensure that they have extremely good threat intelligence to ensure high efficacy. They should also ensure that the vendor providing such solutions has a robust machine learning algorithm that allows the solution to predict attacks. Many of the recursive DNS service providers resolve millions if not billions of Internet requests every day, and they have ML algorithms analyzing the massive amount of data to understand patterns and co-relate patterns by running statistical and machine learning models to identify attacks and thus uncover the attacker's infrastructure.

The secure DNS layer is also easy deployable and doesn't have any requirements on the device itself. It only requires the DNS IP address to be changed from a previous DNS IP address to the secure DNS provider's IP address. Any DNS request coming from the device will now be redirected to the secure DNS vendor's cloud network, which will then resolve all the DNS requests and block any request to the malicious domains.

Enhanced Visibility and Monitoring

One of the most important security capabilities that's required in any organization is enhanced end-to-end monitoring to understand the communication among the devices and between the devices and the network elements, including monitoring the encrypted traffic.

After discussing and deploying proof of value (PoV), which is a marketing term used by many vendors to make solution validation in service provider networks sound cooler, a number of service providers see very little value in aggregating and tapping the user plane data of the devices. In 5G, the user plane data from devices (eMBB slice-related devices) will be in the terabytes of volume. Having a solution for end-to-end user plane (UP) monitoring is not viable due to cost and technical reasons. Control plane, service plane, and OAM are the key layers that should be monitored at minimum. By validating this method in multiple service providers, it is quite clear that many of the anomalies can be detected by monitoring the control plane, service plane, and OAM layer. Once the monitoring for these layers is established, the service provider can pick and choose the UP-layer visibility for specific use cases. IoT devices (related to machine-to-machine use cases), as such, are not user plane intensive, so having granular visibility would not be a major hurdle in terms of cost.

Before investing in an end-to-end monitoring system for the consumer IoT, service providers should try to build a unique ID system, as explained in the section "Identification, Authentication, Access, and Certificate Management" in this chapter. This will also help the service providers in reducing the mean time to repair (MTTR), as the service provider can quickly respond to the unplanned device breakdown.

Figure 8-18 illustrates the monitoring system for anomaly detection for your deployments.

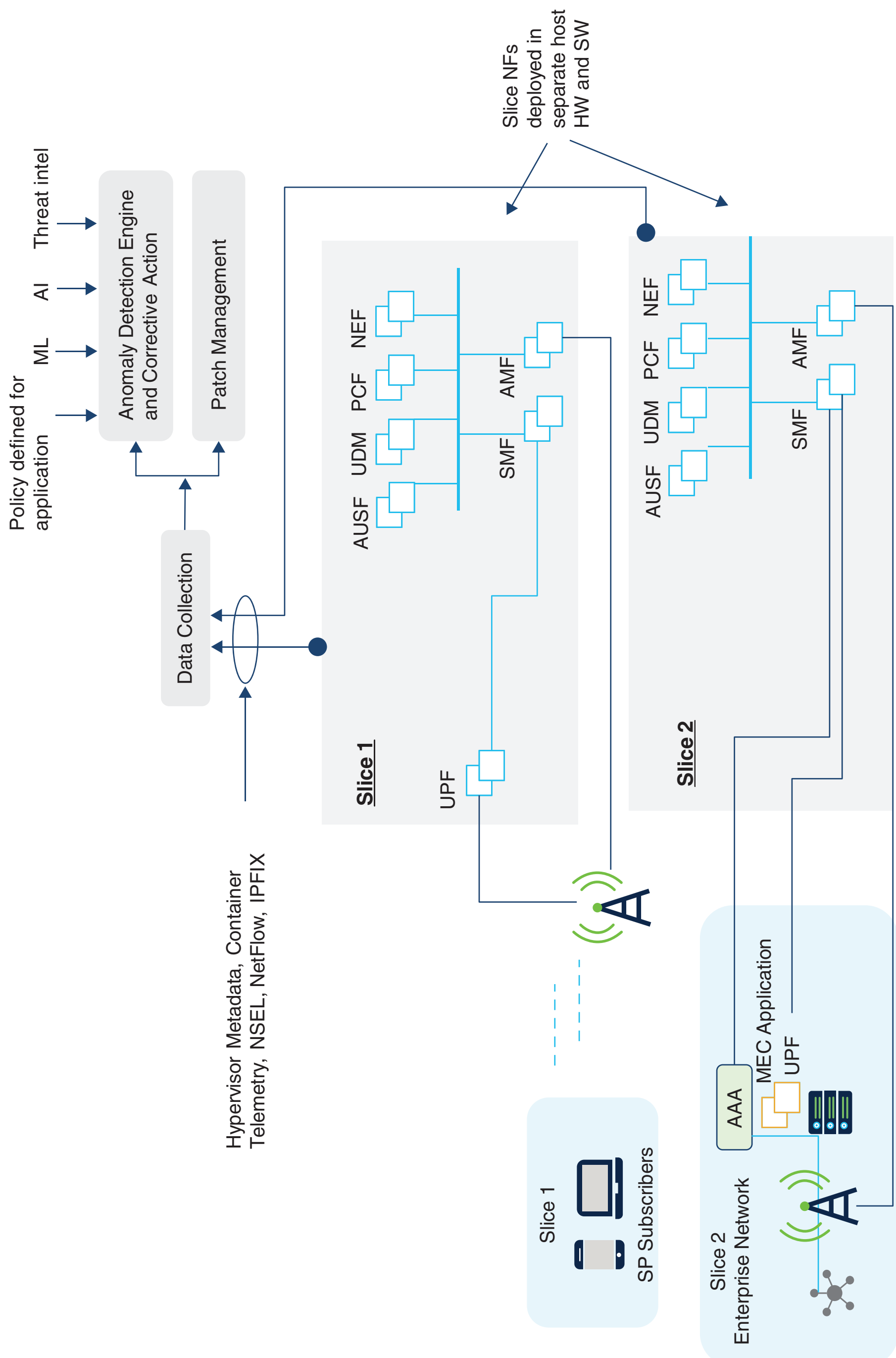

FIGURE 8-18 Enhanced Visibility of IoT Device and IoT Slice Layers

As shown in Figure 8-18, the monitoring solution should also cater for enterprise use cases, as 5G allows easier integration into the enterprise networks using methods such as multi-access and edge computing (MEC) and network slicing. Due to the flexibility in deploying the use cases, the monitoring solution should also follow flexibility and scalability. There are monitoring solutions available in the market today that allow for multivendor packet flow collection (without the need for physical probes) and then analyze the data collected after packet de-duplication and VXLAN striping. Having such monitoring solutions would also support other use cases, such as reusing the same solution for IT and telco DC infrastructure monitoring.

It is also recommended that you look at utilizing monitoring solutions that have integration with the products with capabilities such as responding to any detected anomalies within the device or the device network. The minimum possible response should be the capability to isolate the infected devices or push the devices into a segment that will have access to only critical services.

The visibility and monitoring layer, though very critical, might become very expensive for you if you don't plan it properly for the IoT use cases. One of the methods you could use here to optimize is to consider enhanced visibility and monitoring for control plane, service and management layer of the network functions, and network devices specific to the IoT network. If the IoT network and devices use API-based communications that are encrypted using Transport Layer Security (TLS), it is important to have visibility in the encrypted layer as well. Using a decryption engine and then analyzing the packets, though effective, is not always the best method, as multiple decryption points will reduce the effective security posture of your network. In such cases, it will be more effective to perform malware detection in encrypted traffic without decryption using solutions available today that analyze the encrypted packet header and look at the behavior of cipher suites and so on to determine any anomaly and malicious behavior. Some smart mIoT devices will also provide a basic telemetry with a couple of key counters, which will help you to understand if they have been tampered with. Such IoT devices can be blocked or reported to the IoT device user, depending on the SLA.

Access Control

Access control for 5G SIM or universal integrated circuit card–capable devices are catered for by the inherent 5G Identity and Access Management mechanisms. But many of the consumer IoT devices being deployed for quite some time will use non-3GPP technologies and legacy 3GPP mechanisms and connect to the 5GC using network elements like the non-3GPP Inter-Working Function (N3IWF), which is responsible for the interworking between the untrusted non-3GPP components and the 5GC.

There are various access control mechanisms used by service providers today, primarily role-based access control (RBAC), mandatory access control (MAC), access using security group tags (SGTs), attribute-based access control (ABAC), and so on. For the cloud-hosted IoT management functions such as CM, PM, and FM and provisioning servers catering for consumer IoT devices, a very strict RBAC schema should be applied as a minimum, which is then followed by using multifactor authentication (MFA) for the users and devices. There should be layers of access control for any remote configuration of the IoT subsystem (controller, server, device, and so on).

To ensure that only legitimate users with the right levels of access are accessing the management layer/operational technology (OT) of the IoT network, you should apply zero-trust principles and use mechanisms where you authenticate and re-authenticate the users at varying levels of time and network layers. For example, you should use mechanisms such as MFA, which is integrated into your existing Identity and Access Management (IAM) layer. This integration will ensure that any change in the user's role is mapped to RBAC. If the previous role of the user was admin with privilege access, once the person leaves the organization or changes role, the integration will ensure that the person does not have privileged access anymore. This layer, although foundational, is rarely designed properly due to multiple access control vendors and multiple MFA vendors being deployed at different departments of the service provider. In some cases, there are six to seven multiple IAM solutions deployed in the same domain of the service provider, thus unnecessarily complicating the access control and leading to improper configuration and blind spots.

Figure 8-19 illustrates the granular access control for IoT deployments by providing the secondary authentication mechanism for IoT devices using the enterprise AAA/IAM.

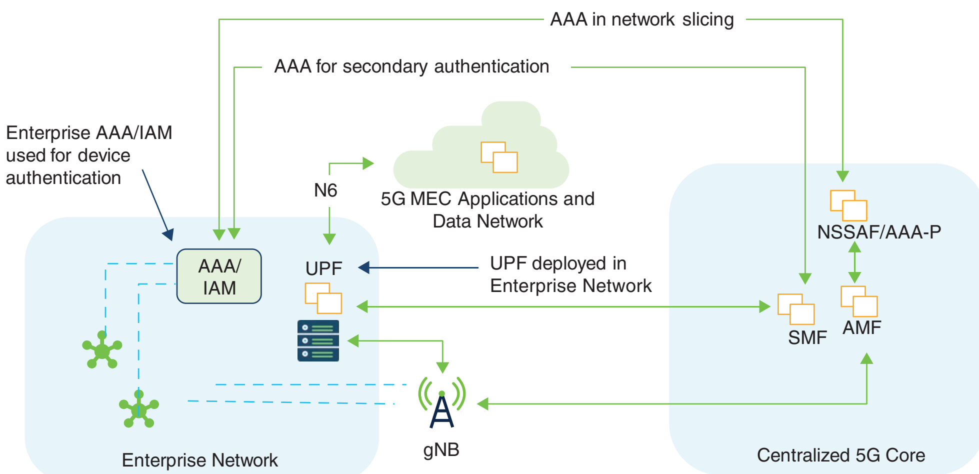

FIGURE 8-19 Granular Access Control for 5G IoT Network

As shown in Figure 8-20, the user will have to go through primary authentication, secondary authentication, secure Internet access, and a granular role-based access for accessing the device and the consumer IoT subsystem.

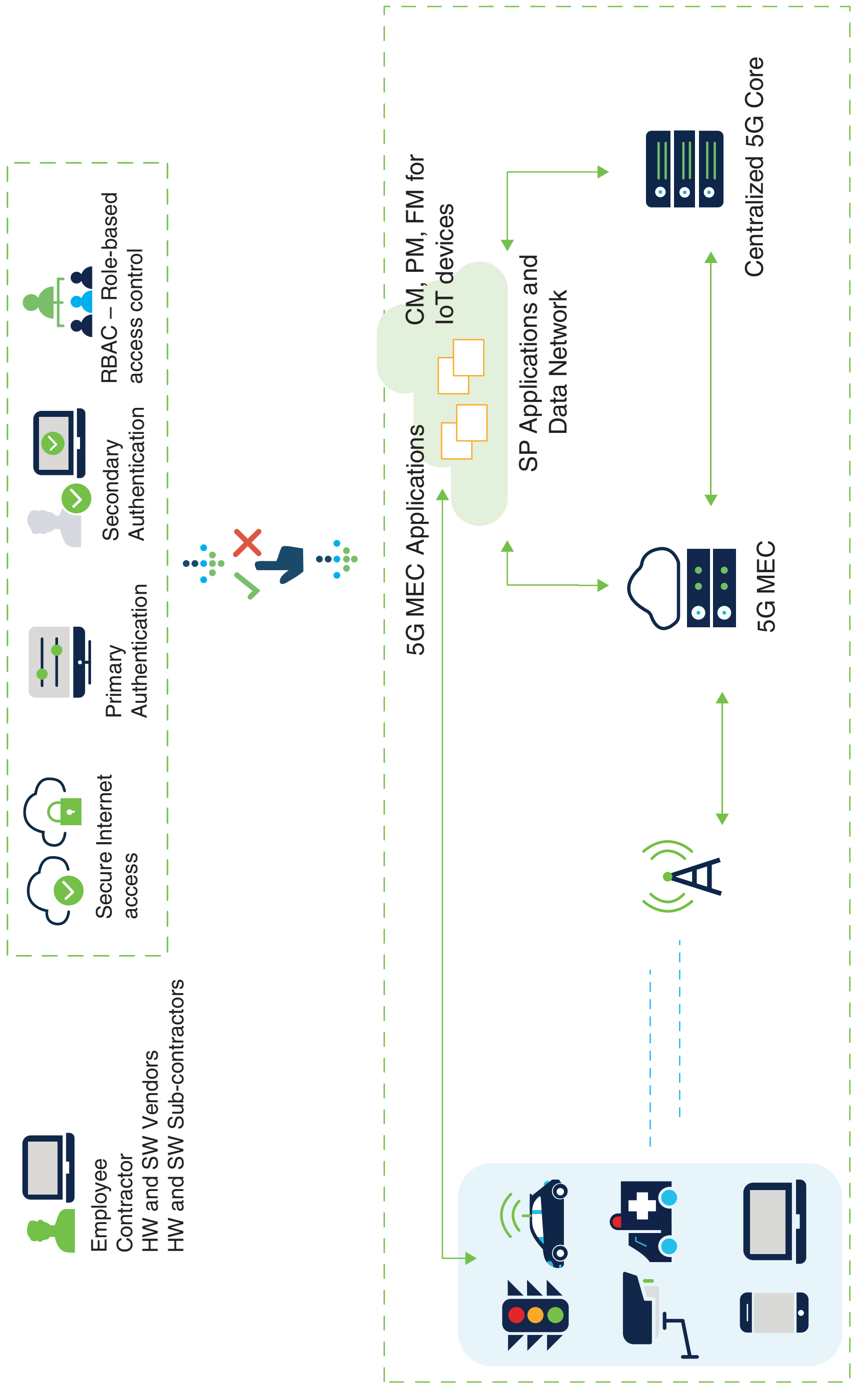

FIGURE 8-20 Granular and Multilayered Access Control for 5G IoT Network and IoT Devices

Index

Numbers

A

B

C

D

E

F

G

H

I

J

K

L

M

O

P

T

V