Whitepaper
Cisco Public

5G Security Innovation with Cisco

Authors: Michael Geller and Pramod Nair

Introduction

5G touches almost every aspect of the way we live our lives. It's not just about faster, bigger or better, it's about utilizing 5G as an enabler to a series of services that we all will consume in every aspect of our lives. The time is NOW to consider the security implications and cyber risk profile that come with 5G. The business operational risk, legal risk and reputational risk of not only the companies who provide 5G transport, but all companies, nation states and individuals who provide the services that will utilize 5G. 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. Many IoT (Internet of Things) services will utilize 5G services. The intersection of 5G and IoT brings an extension of the existing threat surface that requires careful consideration from a cyber risk perspective. This white paper highlights innovative thoughts which enable you to take action and meet the challenges creating a security safety net for the successful deployment and consumption of 5G based services.

5G is as much the application of new architectural concepts to traditional mobile networks as it is about the introduction of a new air interface. The 5G mobile network intentionally sets out to be a variable bandwidth heterogeneous access network, as well as a network intended for flexible deployment. Aside from the usual reasons of generational shifts in mobile networks, i.e. those concerned with the introduction of networking technologies on lower cost curves, the 5th generation of mobile networks has to be able to allow the mobile service providers to evolve towards new business models that may result in future modes of operation that are very different from those of today. This presents a problem from the view point of securing such a network. The need to be flexible increases the threat surface of the network.

Security provides the foundation of service assurance. Adversaries and the threats that they impose against the networks used to deliver critical services continue to get smarter, more agile, and more destructive. Networks used to deliver applications continue to converge, making it more important to properly segment threats and vulnerabilities by domain, while examining the aggregate threat landscape at the same time. Examples of this include the evolved packet core where traditional and mobile services share an infrastructure leveraging the carrier data center and cloud for operational efficiency and also for service delivery. Cisco's architectural innovations and evolution of existing networks to meet the needs of new service models like IoT services pushing technology evolution such as mobile edge compute and widely distributed secured data centers introducing a new set of visibility and control elements to handle the evolved threats.

In order to properly secure the "full stack" that delivers a connected application, two fundamental elements are applied: visibility and control. Visibility refers to the ability to see and correlate information from the carrier cloud to baseline proper behavior and then to measure deviation from that norm. Simply said, "If you can't measure it, you can't manage it." Sources of visibility come from traditional network measurements (netflow, open flow, etc.), but the need to measure all aspects of a flow, from all elements of the carrier cloud to the application to the end customer, has changed what data is collected and where we get it. An example of the new visibility includes the use of application level probes that are synthetically generated and travel through the network to get a clear picture of how an application is behaving. Another example is where the Path Computation Element, which has a near real time database representing the network topology, is queried programmatically to determine the impact of a potential mitigation action on critical service classes for DDoS. Once all of the telemetry is gathered, a security controller and workflow will analyze it and determine, based on policy, suggested mitigation and controls to be applied. Of course, we have an iterative loop of constant learning. The Cisco Talos research team keeps our customers ahead of the game by its threat research and deployment of mitigation rules into our full portfolio of products, removing that burden from the Service Provider allowing them to focus on their core competencies.

Control refers to the actions taken to mitigate an attack. Some controls are taken proactively while others are applied after an attack takes place. There are two types of attacks. Day zero attacks are threats that we don't previously have a fingerprint for. Typically deviations in known good behavior of the carrier cloud and applications that request service and state from it, are identified by the security controller and some action is then taken to mitigate the attack or to get additional visibility, an action sometimes taken to properly identify the adversary. Day one attacks are threats that we have a signature or fingerprint for and, quite often, a mitigation strategy exist in advance to handle the attack. Controls take the form of modifications to the carrier cloud to apply quality of service changes in per hop behavior to minimize the impact of an attack and also take the form of physical and virtual security assets applied as close to the source of the threat as possible in order to minimize collateral damage.

The information that the operator has that delivers the application is vast. Innovation in the way that we apply the information we have, in a close loop iterative process, is a recent innovation in threat visibility and mitigation. This is where automation, orchestration and NFV meets security to solve today and tomorrow's security needs. The three elements of the closed loop iterative process are: policy, analytics, and the application delivery cloud (the whole transaction from the application to the networks used to serve it). Operators can now apply innovative methods to correlate geo-location information to behavioral analytics, compare those against policy in the context of a threat to the carrier cloud, and ascertain the nature of that threat and what to do about it with far greater clarity. Visibility and control properly applied to the advanced threats of today offer the carrier cloud a level of protection. We must continue to evolve, grow and get smarter to keep our networks safe and resilient in the time of attack.

In the end, security is about finding threats faster, fixing them faster and learning all of the time.

What's New With 5G?

5G is the next generation of 3GPP technology, after 4G/LTE, being defined for wireless mobile data communication. Starting 3GPP Release 15 onwards 3GPP has started defining standards for 5G. As part of 3GPP Rel 15 New 5G Radio and Packet Core evolution is being defined to cater to the needs of 5G networks. [1] & [2] provide more details on 3GPP standards for 5G architecture.

Below listed are the some of the key end goals of 5G:

- Very high throughput (1-20 Gbps).
- Ultra low latency (<1ms).
- 1000x bandwidth per unit area.
- Massive connectivity.
- High availability.
- Dense coverage.
- Low energy consumption and
- Up to a 10-year battery life for machine type communications.

There is a security focused "expert teams" that is a part of many organizations driving 5G architectures. 3GPP and NGMN are two such organizations. This is empowering key 5G security topics into the broader 5G architecture evolution. These topics include authentication, encryption, placement of security controls and sources of visibility. This is all driven by a set of new use cases which drive the 5G architecture.

Below are some of the projections set by 5G PP (A Joint initiative between EU Commission and European ICT):

5G PPP will drive the future networked society

Increasing wireless capacity **1,000 times**

Connecting **7 billion** people

Connecting **7 Trillion** "things"

Saving **90%** energy

Perceiving **zero** downtime

Source: 5G PPP

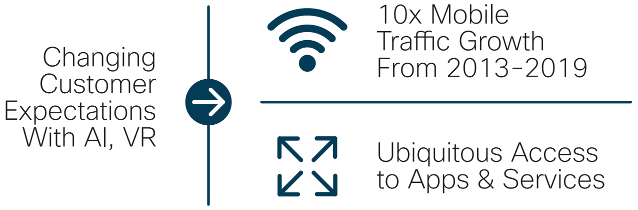
5G will bridge wireless and wireline networks, forcing a major network architectural change from radio access to core.

Cisco’s 5G Vision

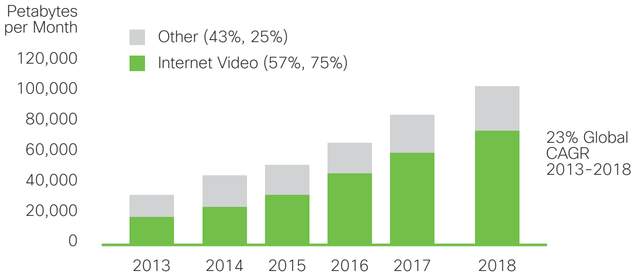
5G is an enabler for new set of possibilities and capabilities. Every new generation of 3GPP wireless mobile data communication technology has set the stage for new set of use cases and capabilities. 3G was the first truly wireless mobile data communication technology that catered to data communication. Whereas 4G was the first truly all IP wireless data communication technology. Both 3G and 4G have been instrumental and foundational to the data communication over mobile devices which led to proliferation of applications like video, ecommerce, social networks, games and several other applications on mobile devices. Focus in 3G/4G was more on mobile broadband for consumers and enterprises.

Below are some trends and new opportunities that the operators have in front of them:

The world has gone mobile


Changing Customer Expectations With AI, VR
10x Mobile Traffic Growth From 2013-2019
Ubiquitous Access to Apps & Services

Traffic growth in Access, driven by video



Rise of cloud computing

Soon to Change SP Architectures/ Service Delivery
Changing Enterprise Business Models Efficiency & Capacity

Digitization leading to IoE

Emergence of the Internet of Everything

At the same time new sets of use cases are being introduced that is going to throw up new sets of challenges, complexities and threats. Thus, new 5G network has to help operators manage current needs as well as gear up for new needs of the upcoming new use cases. 5G is not just going to be about high-speed data connections for enhanced mobile broadband but will enable several new capabilities that can cater to several new enterprise use cases. Securing the “enterprise network slice” presents a number of new challenges required to securely deliver the outcomes that enterprises who use 5G require, both operationally and by regulatory control. 5G will not just be about serving consumer and enterprise subscribers with high throughput connectivity. 5G will enable new revenue avenues and opportunities for operators by being able to cater to requirements for several new enterprise use cases. Thus, Cisco envisions 5G to equip operators with more capabilities to cater to enterprise customer needs to support their current as well as new use cases. The delivery of these new use cases is predicated on a “safety net” provided by security visibility and controls pervasively throughout the 5G network, from transport all the way up to the applications. 5G provides a number of new threat boundaries and to make it even more difficult, those boundaries are transient meaning that they can move.

Cisco views 5G core as an opportunity for service providers to take advantage of the major changes taking place in the data center, networking and the economics of mobility in a standardized multivendor environment. Very significant changes are being defined for the mobile core that facilitate new opportunities such as personalized networks through slicing and more granular functions. 5G provides a framework to take advantage of the massive throughput and low latency that new radio provides.

Below listed are some of the use cases that 5G will cater to:

Transportation
Autonomous Vehicles
Automotive

IoT

Augmented Reality
Virtual Reality

Smart City
Traffic Management
Emergency Services

Manufacturing
Robotics

Tactile Internet

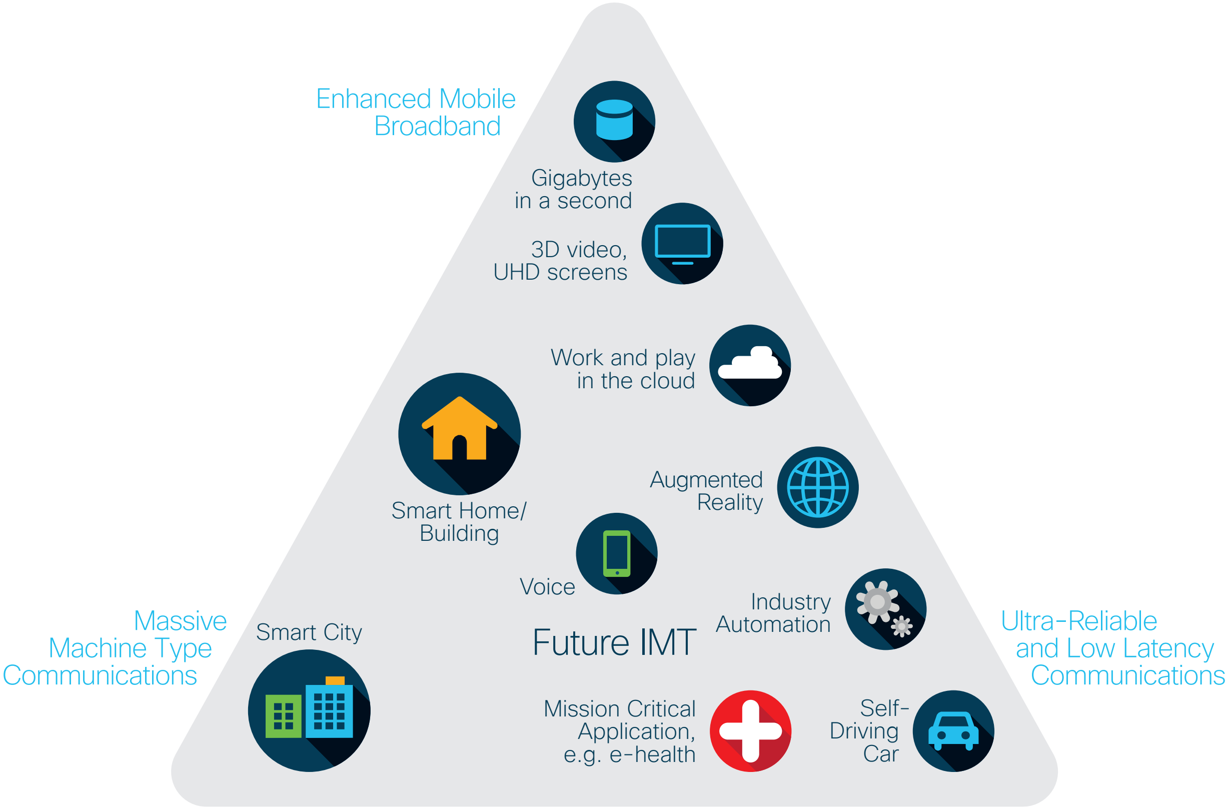
Health
Fitness & Healthcare

Smart Grid
Utilities

Government

Smart Office

Below is the chart that illustrates the broad categories of use cases (source: ITU) that 5G will cater to:


Enhanced Mobile Broadband
Gigabytes in a second
3D video, UHD screens
Work and play in the cloud
Smart Home/ Building
Augmented Reality
Voice
Industry Automation
Massive Machine Type Communications
Smart City
Future IMT
Ultra-Reliable and Low Latency Communications
Mission Critical Application, e.g. e-health
Self-Driving Car

Source: ITU

Below listed are the 3 different Use case categories that will cover all the use cases.

Enhanced Mobile broadband (eMBB): 5G Enhanced Mobile Broadband (eMBB) brings the promise of high speed and dense broadband to the subscriber. With gigabit speeds, 5G provides an alternative to traditional fixed line services. Fixed wireless access based on mmWave radio technologies enables the density to support high bandwidth services such as video over a 5G wireless connection. To support eMBB use cases, the mobile core must support the performance density, scalability and security required.

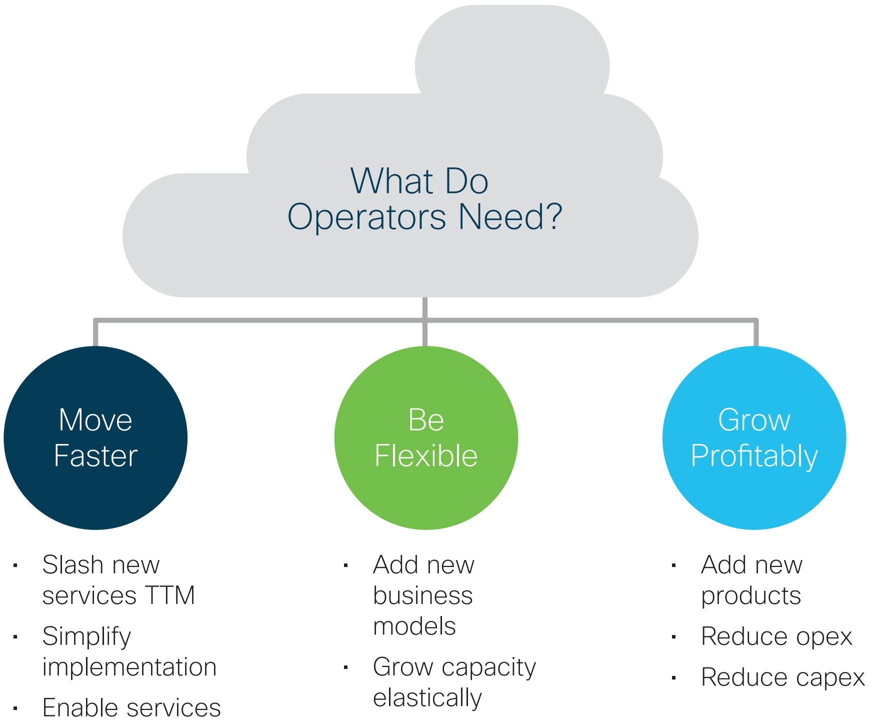
Ultra-reliable low latency Communications (Robotics, Factory Automation): Ultra-reliable low latency communications (URLLC) focuses on mission critical services such as augment and virtual reality, tele-surgery and healthcare, intelligent transportation and industry automation. Traditionally over a wired connection, 5G offers a wireless equivalent to these extremely sensitive use cases. URLLC often requires the mobile core User Plane Function (UPF) to be located geographically closer to then end user in a Control and User plane Separation (CUPS) architecture to achieve the latency requirements.

Massive IOT: Massive IOT in 5G addresses the need to support billions of connections with a range of different services. IOT services range from devices sensors requiring relatively low bandwidth to connected cars which require a similar service to a mobile handset. Network slicing provides a way for service providers to enable Network as a Service (NaaS) to enterprises; giving them the flexibility to manage their own devices and services on the 5G network.

Below are characteristics of these use cases:

- Efficient low-cost communication with deep coverage.
- Light weight device initialization and configuration.
- Efficient support of infrequent small data for mobile originated data only communication scenarios.

Given the several different use cases, each with different needs and requirements, the network complexity and costs for operators will increase. Thus, in order for operators to be able to cater to these different use cases and stay ahead of the competitors, Cisco believes operators will need following capabilities:



What Do Operators Need?

Move Faster
- Slash new services TTM
- Simplify implementation
- Enable services

Be Flexible
- Add new business models
- Grow capacity elastically

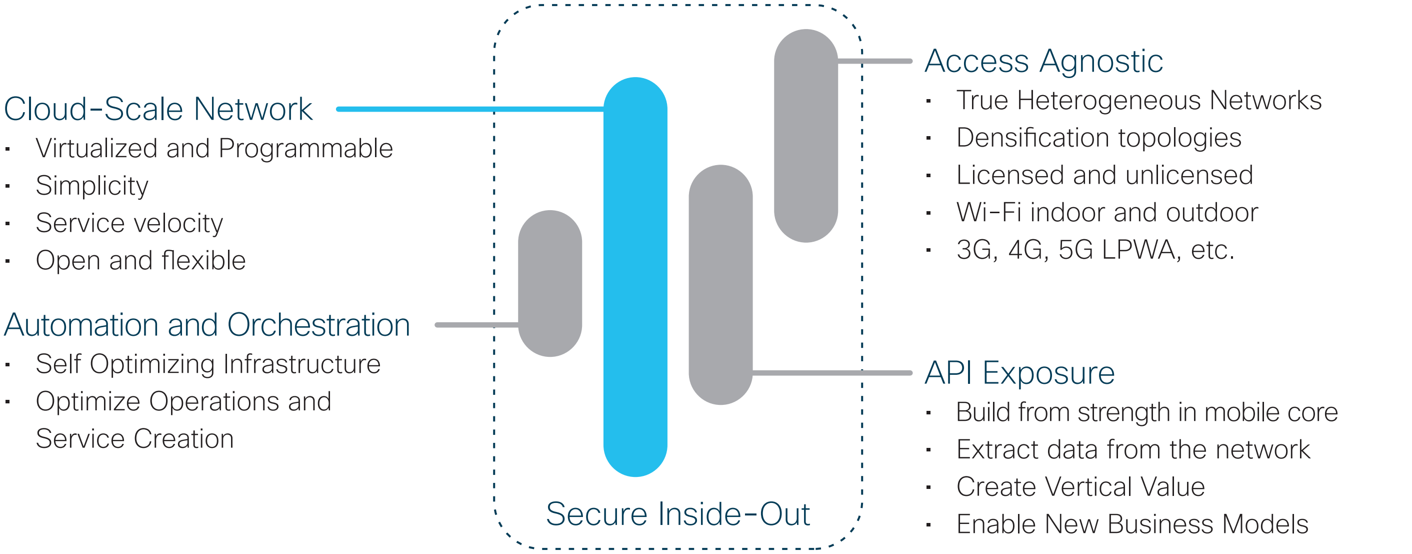
Grow Profitably
- Add new products
- Reduce opex
- Reduce capex

Cisco is developing its 5G solution keeping above operator needs in mind. Cisco's strategy is to transition its customers to cloud centric world to be able to get the benefits of cloud native solution and thus equip them to be able to meet the needs listed above. Cisco believes 5G is not just about new radio, but it is the total end-to-end network including both RAN and packet core need to evolve to cater to these needs of operators.

One very interesting evolution in securing the operator's network is the role of encryption. Most security controls require visibility to then make a decision against policy at the security control used to mitigate the threat as close to the source as possible to minimize the collateral damage of the threat. About half of all traffic on the Internet is encrypted today and it is expected to increase over the next few years. Cisco's innovation in encrypted traffic analytics provides a machine learning based set of technology to deal with this evolutionary aspect of 5G.

The widely distributed data center brings with it a greatly expanded threat surface. Many attackers will leverage advanced persistent threats which infiltrate the target network and move laterally to cause damage and steal data. Segmentation must be applied to make it very difficult for a criminal or hacker to move laterally, protecting the network and applications using the structure of the network. The Cisco Data Center leverages automation delivered by an identity controller and delivered programmatically to all network elements where security must by dynamically provided. That way, a simple matrix can be used to dynamically apply the policy and to avoid any errors associated with manual application of policy. Segmentation applies throughout the 5G network, in the transport, the distributed data center delivering mobile edge compute and other key services, and in other parts of the 5G architecture too. Segmentation is delivered using two technologies: Segment routing in the transport and TrustSec in the data center. Cisco is uniquely able to deliver end to end segmentation.

Below listed are key tenets of the Cisco 5G Solution Architecture:



Cloud-Scale Network

- Virtualized and Programmable
- Simplicity
- Service velocity
- Open and flexible

Automation and Orchestration

- Self Optimizing Infrastructure
- Optimize Operations and Service Creation

Secure Inside-Out

Access Agnostic

- True Heterogeneous Networks
- Densification topologies
- Licensed and unlicensed
- Wi-Fi indoor and outdoor
- 3G, 4G, 5G LPWA, etc.

API Exposure

- Build from strength in mobile core
- Extract data from the network
- Create Vertical Value
- Enable New Business Models

Cisco is a leading Packet Core vendor for decades and has been influencing 3GPP standards given its expertise it has built over several years. Cisco has witnessed transitions earlier too, firstly from 2G to 3G and then 3G to 4G and is currently best placed vendor to define and lead the solution for the important and crucial transition from 4G to 5G.

Platform based VPC solution is deployed in 40+ networks globally making Cisco one of the leading virtual packet core vendor.

Cisco has been working on several packet core concepts even before they could get standardized in 3GPP. Security implications of each of these will be addressed below. For instance, Cisco was one of the vendor to demonstrate CUPS in MWC in 2016 and 2017 even before 3GPP could standardize it. Continuing the similar trend Cisco is aggressively working in introducing a pre- standards version of the 5G solution and thus evaluate needs the needs of the NextGen 5G network and take them to 3GPP and influence the standards accordingly.

Below listed are some of the reasons for Operators to choose Cisco 5G solution:

Virtualization	Open	R&D Investment
Market leader with 40+ deployments including majority T1 operator	Market leader with 40+ deployments including majority t1 operator	Significant R&D investment in cloud native and NFV technologies
VNF Portfolio	**Full Stack**	**Virtualization**
Comprehensive “off-the-shelf” VNF catalog	Full stack capabilities (over cloud, under cloud)	Align Cloud Native technology with 5G architectural vision of automated/distributed/edge based intelligent network

3GPP has defined two different solutions for 5G networks: 5G Non-Standalone (NSA) and 5G Standalone.

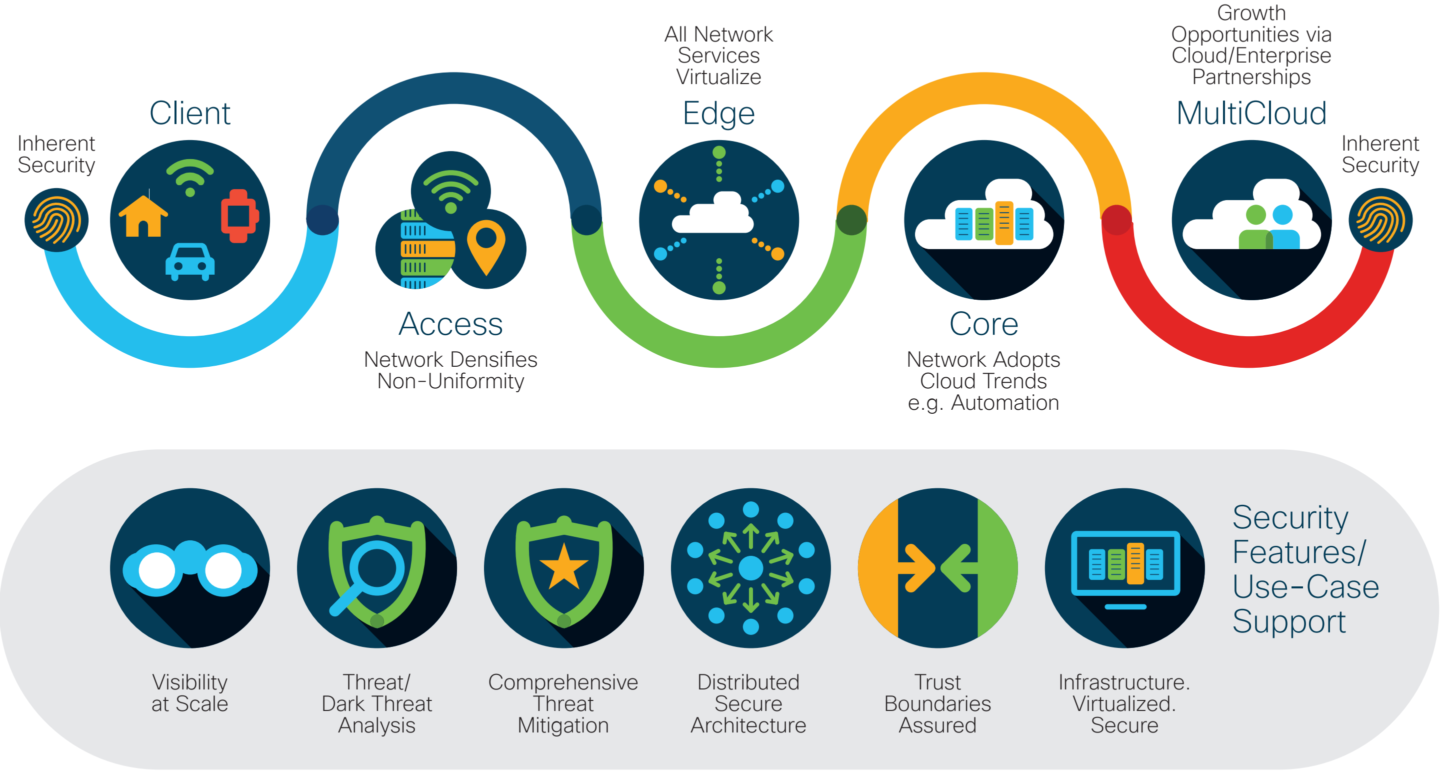
5G Non-Standalone Solution (NSA): In 5G NSA Operators will leverage their existing EPC packet core to anchor the 5G NR using 3GPP Release 12 Dual Connectivity feature. This will help operators with aggressive 5G launch needs to launch 5G in shorter time and with lesser cost. 5G NSA solution might suffice for some initial Use cases. But 5G NSA has some limitations with regards to getting a much cleaner truly 5G Native solution and thus all the Operators will eventually be expected to migrate to 5G Standalone solution.

5G Standalone (SA) Solution: In 5G SA an all new 5G packet core is being introduced. It is a much cleaner with several new capabilities built inherently into it. Network Slicing, CUPS, virtualization, automation, multi-Gbps support, ultra low latency and other such aspects are natively built into 5G SA Packet Core architecture.

Cisco will have in its portfolio packet core solution for both 5G Non-Standalone (NSA) and 5G Standalone (SA) network. Cisco's goal is to come up with 5G Packet Core solution that allows operators to make transition from 4G to 5G in a graceful step-by-step manner.

Security Innovation and Thought Leadership for 5G

Cisco 5G - Security Roadmaps Developing in Parallel


Inherent Security
Client
Access
Network Densifies Non-Uniformity
All Network Services Virtualize
Edge
Core
Network Adopts Cloud Trends e.g. Automation
Growth Opportunities via Cloud/Enterprise Partnerships
MultiCloud
Inherent Security
Visibility at Scale
Threat/ Dark Threat Analysis
Comprehensive Threat Mitigation
Distributed Secure Architecture
Trust Boundaries Assured
Infrastructure. Virtualized. Secure
Security Features/ Use-Case Support

As the graphic above shows, Security is a foundational aspect of Cisco's 5G strategy. There are five aspects of the foundational security fabric which deliver secure outcomes in the Cisco 5G network. Some are evolutionary. Others are revolutionary. In aggregate, the threat surface of today and tomorrow is addressed providing the operator and consumer a level of service assurance for critical 5G based services.

The five focus areas for security innovation and thought leadership are:

1. The architecture and trust boundaries detailing the threat surface (now and tomorrow) of 5G and IoT
    a. Where the Enterprise meets the 5G slice
    b. Where SP IT meets 5G
2. Technology trends and architectures impacting how the 5G network is secured
3. Visibility at Scale
4. Threat and Dark Threat Analysis
5. Comprehensive Threat Mitigation

These five areas are covered in the remainder of the white paper below.

The graphic below adapts these five areas to 5G operational security requirements:

Stop threats at the edge

Protect users wherever they work

Control who gets onto your network

Simplify network segmentation

Find and contain problems fast

The Architecture, Technology Trends and Trust Boundaries – The Threat Surface (now and tomorrow) of 5G and IoT

5G and Evolving Architectures

5G is adopting new and adapting existing networking concepts into its architecture. Some of these architectural shifts are, in effect, a modernization of the mobile architecture to fit within cloud operation. An example of this shift is the separation of control plane and user planes (CUPS). While other shifts are due to an appreciation that some new use cases. An example of this shift would be augmented and virtual reality (AR/VR), which may require localization and/or are latency sensitive processing that the network edge (MEC). In this section, we briefly outline the various 5G architectural enablers to 5G. Naturally, not all of these architectural concepts need be applied simultaneously to a 5G network and in reality, may be introduced as pragmatism and opportunity dictate, therefore Cisco expects that there would be more than one architecture which would be adhered to by various operators. This influences and impacts the overall security architecture that is adopted by the service provider.

Some of the evolving architecture enablers being considered by different operators are:

CUPS

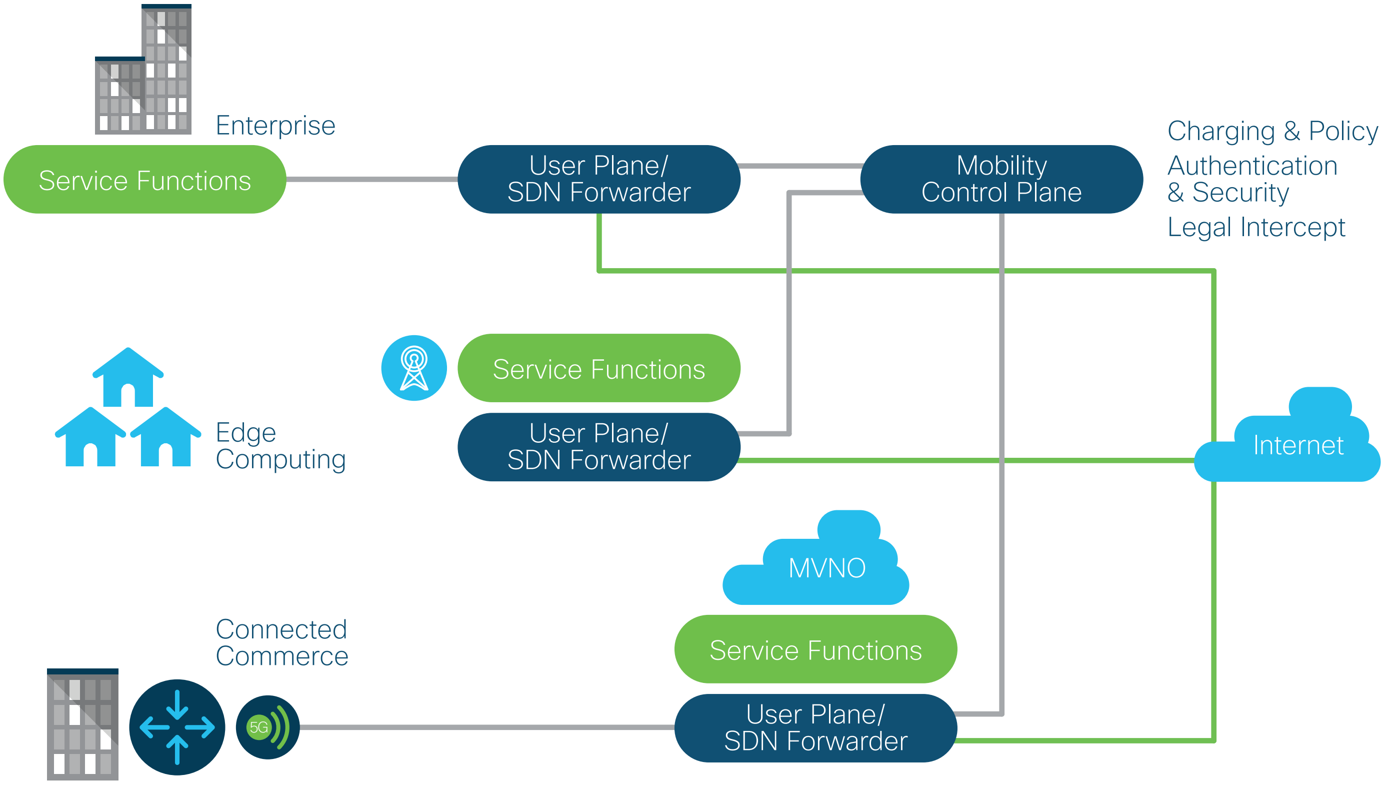


Figure 1. High level CUPS Architecture

Control Plane User Plane Separation (CUPS) architecture (see Figure 1) allows the distribution of the core elements, leading to greater function utilization efficiencies and the placing the network functions into the network where most appropriate given service constraints of an SLA for example. CUPS also allow efficient network function scalability based on the traffic dimensioning.

MEC

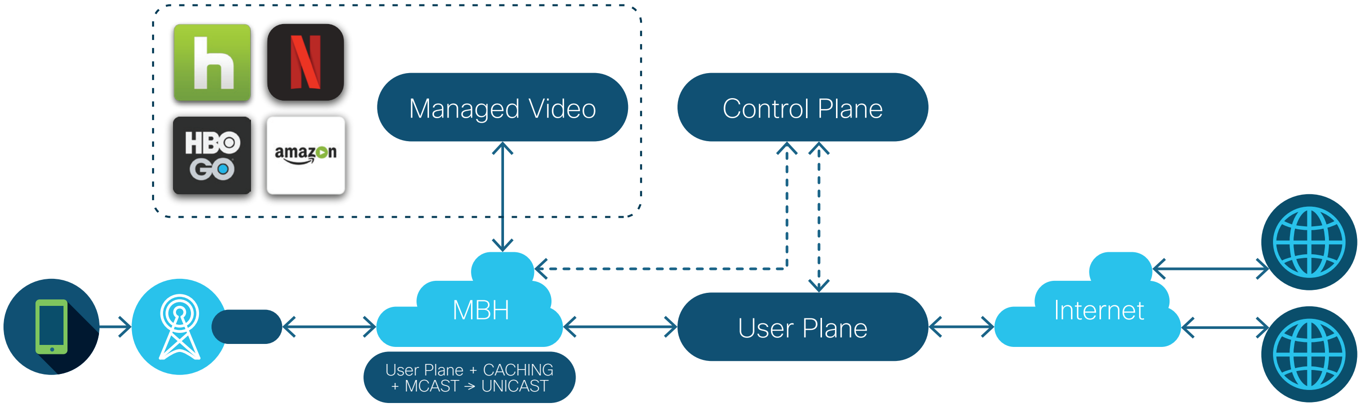


Figure 2. High level MEC Architecture

Mobile Edge Computing (MEC) reuses the CUPS architecture to allow the user plane functions and applications to be placed closer to the network edge (see Figure 2). Moreover, the MEC architecture potentially allows low latency use cases to be fulfilled, since the user applications and services can be located closer to the users. Network traffic management mechanisms, such as offloading etc., can also benefit from the MEC based architecture.

CRAN

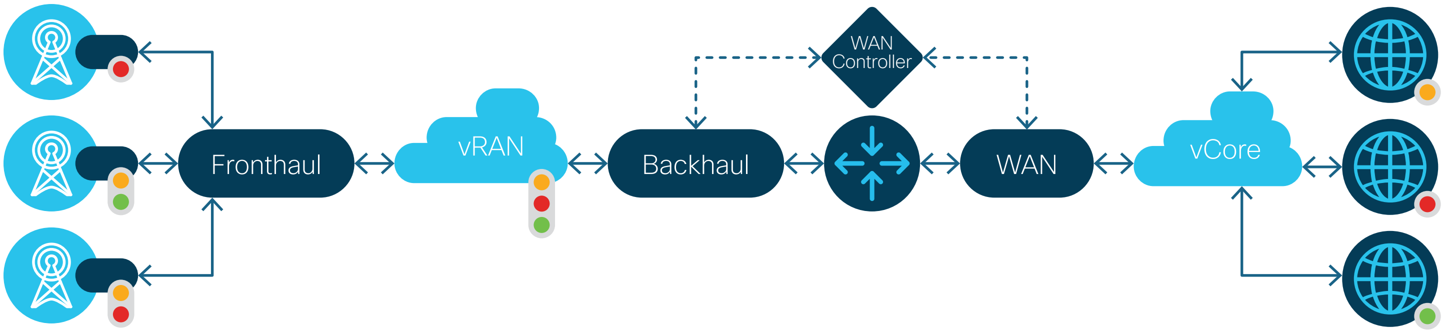


Figure 3. High level CRAN Architecture

Cloud Radio Access Network (CRAN) architecture uses a form of control and user plane separation for the access network. This leads to a split access architecture, whereby some of the RAN processing is done virtually in a "central" edge cloud location, with the remainder of the processing being done in a remote "physical" location such as a remote radio head (RRH). This promotes a front haul and backhaul split in the transport network (see Figure 3). The CRAN architecture can also be combined with network slicing, cloud native Core to allow more flexible 5G use cases.

Network Slicing

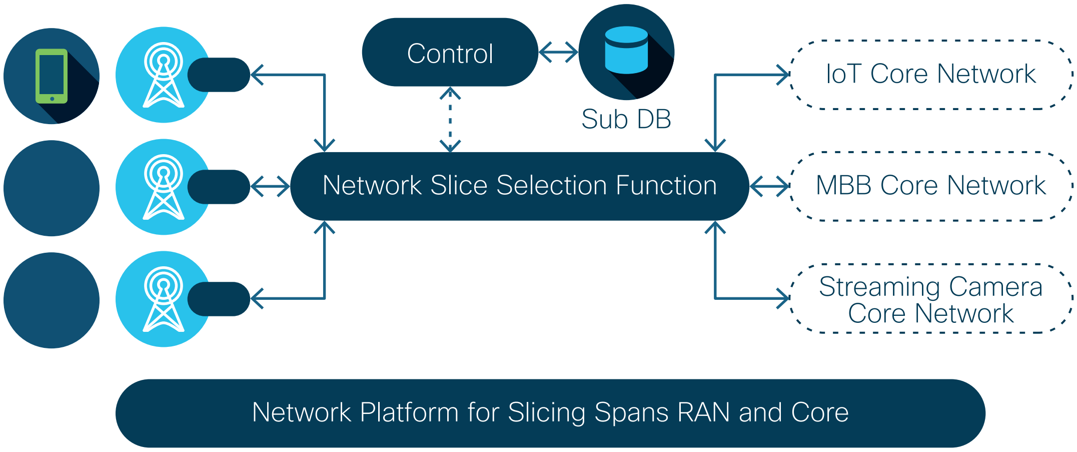

Control
Sub DB
IoT Core Network
Network Slice Selection Function
MBB Core Network
Streaming Camera Core Network
Network Platform for Slicing Spans RAN and Core

Figure 4. High level Network Slice Architecture

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 (see Figure 4). Network slicing, although used today in sparse manner for Enterprise use cases, is expected to be a fundamental architecture component of the 5G network, fulfilling the majority of the 5G use cases.

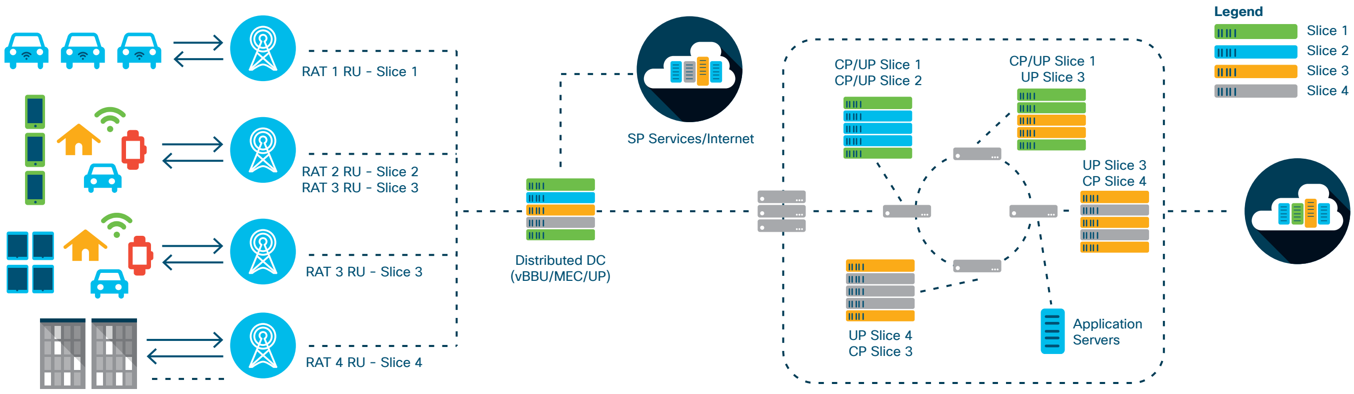
Many operators are considering the offer of a network slice per enterprise, not that dissimilar to the per APN offer for an enterprise in play today. As we consider the points where the enterprise then touches the 5G slice, a number of security aspects must be addressed. Cisco has a dominant role in enterprise security today and can extend that footprint into the 5G slice. These areas include:

Identity, Network Access Control and Segmentation	Cisco Identity Services Engine in the Enterprise
Unified transport for sharing identity	Use of pxGrid to deliver/share identity services consistently
Seamless transition of Enterprise to 5G connectivity	WiFi to 5G handoff with IP Video/Voice
Protect the transient Enterprise edge	Extend virtual and cloud security services footprint
Protect the application boundary	Extend API security and Data Loss Prevention
Segment for threat agility and handling	Adding the 5G slice to the segmented enterprise network

When the architectural concepts outlined in the previous section are combined into the 5G architecture, the threat surface for 5G expands. The threat surface expands for a number of main reasons, among which are:

- The physical structure of the network is changing, (future) applications and use cases need to compute and storage locations closer to the edge for reasons of localization and latency – in effect this is a new public execution site that was not available before, and
- The structure of the networking functions has changed from physical to virtual implementations, and the functions virtualized components can be placed across distributed edge and centralized core clouds.
- There is an emphasis on flexible software based architecture enablers such as SDN (Software Defined Networks), SDA (Software Defined Access), SDR (Software Defined Radio).

In short, we could summarize this situation in the following way, the majority of the threat surfaces in the 5G is due to the network architecture being more flexible and open towards the internet. Figure 5 shows the 5G Architecture which encompasses the characteristics above (through CUPS, MEC, CRAN and slicing) from a threat perspective.



Device Threats
Malware
Sensor Susceptibility
TFTP MitM attacks
Bots DDoS
Firmware Hacks
Device Tampering

Air Interface Threats
MitM attack
Jamming

RAN Threats
MEC Server Vulnerability
Rogue Nodes

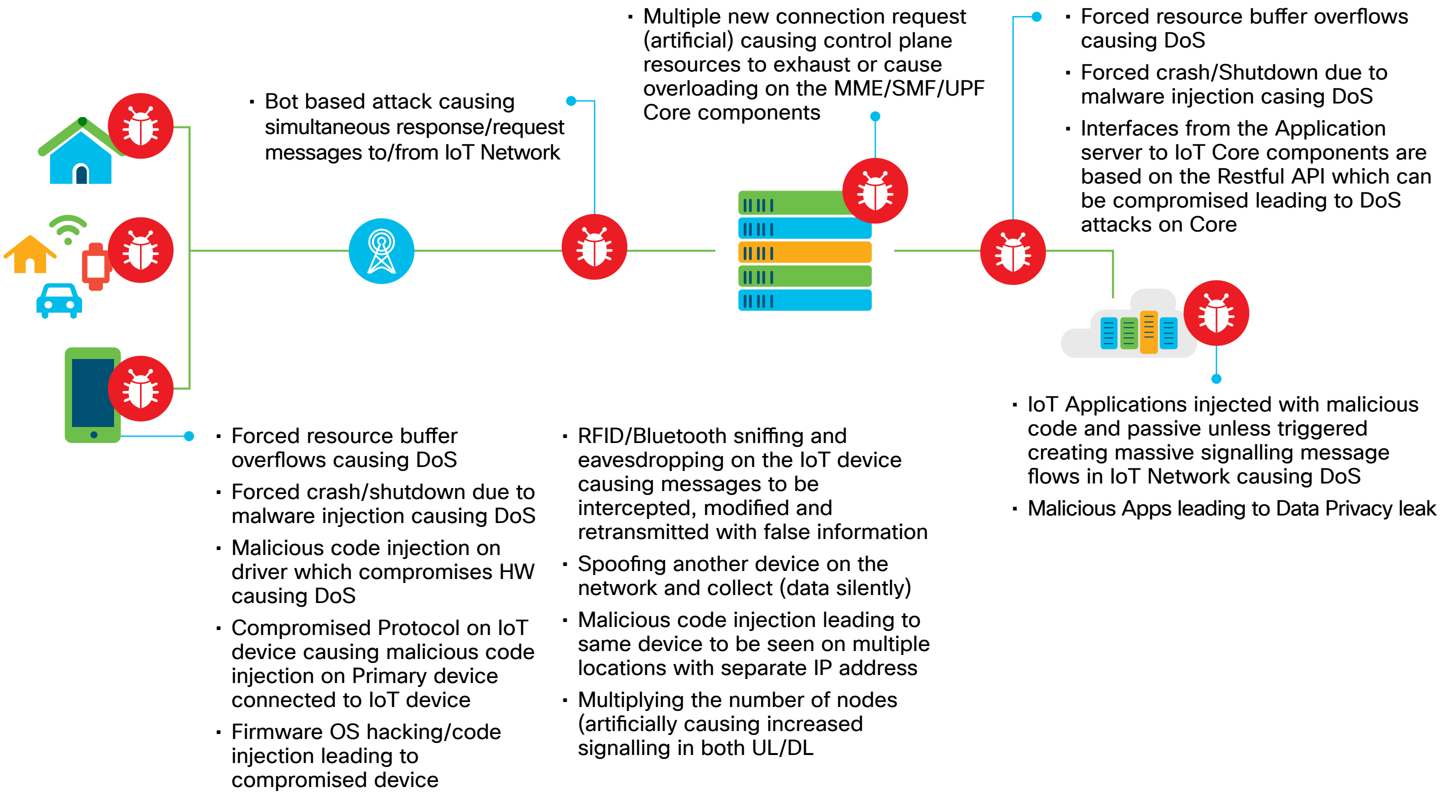
Backhaul Threats
DDoS attacks
CP/UP Sniffing
MEC Backhaul sniff

5G Packet Core & OAM Threats
Virtualization
Network Slice security
API vulnerabilities
IoT Core integration
Roaming Partner vulnerabilities
DDoS & DoS attacks
Improper Access Control

SGI/N6 & External Roaming Threats
IoT Core integration
VAS integration
App server vulnerabilities
Application vulnerabilities
API vulnerabilities

Figure 5. The 5G Architecture Threat Surface.

As one may expect, once we begin to apply specific use cases to the 5G architecture we can analyze what the threat surface looks like for that use case. Figure 6 is an example of this and shows the end-to-end threat surface for a 5G IoT network architecture.



- Bot based attack causing simultaneous response/request messages to/from IoT Network

- Multiple new connection request (artificial) causing control plane resources to exhaust or cause overloading on the MME/SMF/UPF Core components

- Forced resource buffer overflows causing DoS
- Forced crash/Shutdown due to malware injection casing DoS
- Interfaces from the Application server to IoT Core components are based on the Restful API which can be compromised leading to DoS attacks on Core

- Forced resource buffer overflows causing DoS
- Forced crash/shutdown due to malware injection causing DoS
- Malicious code injection on driver which compromises HW causing DoS
- Compromised Protocol on IoT device causing malicious code injection on Primary device connected to IoT device
- Firmware OS hacking/code injection leading to compromised device

- 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 collect (data silently)
- Malicious code injection leading to same device to be seen on multiple locations with separate IP address
- Multiplying the number of nodes (artificially causing increased signalling in both UL/DL

- IoT Applications injected with malicious code and passive unless triggered creating massive signalling message flows in IoT Network causing DoS
- Malicious Apps leading to Data Privacy leak

Figure 6. Example - 5G IoT Threat Surface

The above two figures clearly articulate the threat surface between the different components and explained in detail below.

Various use cases such as M2M, Industry automation, IoT devices use proprietary radio access technologies, or 3GPP and Non-3GPP technologies such as WLAN along with new 5G radio access allows better efficiency for device power consumption and low bandwidth requirement for low cost devices. The low devices do not have the security features such as built in trust anchor which would mitigate threats like firmware and OS hacks. In the majority of use cases with such devices the access is provided by the Mobile Network operators by using credentials and AAA managed by the enterprise. Such devices are prone to Man in the Middle attacks (MitM), firmware and OS hacks, snooping and sniffing attacks, Botnet type attacks where the IoT devices start signaling overloads with the Mobile Network components of the operator.

In 5G there is the potential to move functionality and applications around the network. For example, in Mobile Edge Computing (MEC), the User Plane (UP) and cellular functionalities would move to the edge of the network. In this new architecture, IP connectivity would terminate at the edge of the operator network such as the N6 interface defined in 3GPP. It is likely that the current functionality of DNS resolution and content delivery networks would also move closer to the edge of the network. This situation will alleviate many challenges faced for use cases which require low latency. Examples of such applications are: applications for connected cars when optimizing encrypted video content encrypted end to end (UE to Video Server), since the content would now be delivered from replicas in the operator network. This leads to a new set of threat vectors for the mobile network operator.

Aside from traditional attacks against servers and caches (e.g. via HTTP response splitting), new threat vectors arise. For instance, Denial of Service (DoS) attacks can cause major disruption to the latencies service level agreements committed to by operators. In this scenario, since a very large number of caches at the edge of the network would be deployed to cater for large number of subscribers using low latency applications (ex: video caching), attackers will be able to easily overwhelm these caches with request for content not likely to be used by non-malicious users. This situation would result in filling local caches with "useless" content unusable by subscribers. The vulnerabilities which might cause this to happen would be through traditional attacks on hardware components of the infrastructure, application vulnerabilities, API which are not properly secured and rogue nodes within the architecture.

The 5G core architecture also creates entirely new security threat vectors arising out of virtualized mobile network components and separate slices being created for 5G use cases, and the exposure of the mobile network core components towards the 3rd party applications and external internet facing interfaces using exposure functions specified by 3GPP such as SCEF (Service Capability Exposure Function) and NEF (Network Exposure Functions). The various architecture evolutions towards 5G also require interconnections with existing 3GPP infrastructure such as Mobility Management Entity (MME), Serving Gateway (SGW), Packet Gateway (PGW), and the subscriber authentication and authorization infrastructure such as AAA & Home Subscriber Server (HSS) for 3GPP & non-3GPP access being used in LTE. This increases the treat surface in diameter, Stream Control Transmission Protocol (SCTP) & GPRS Tunneling Protocol (GTP) protocols which we see today along with the new threat surfaces such as DoS & DDoS attacks on network slices due to artificial resource exhaustion, side channel attacks across slices due to class attacks on implementations of cryptography, impersonation attacks against Network Slice managers and the Orchestration layer, control plane signaling flood in the Control Plane User Plane Separation (CUPS) architecture, securing API and Applications which communicate with the SCEF & the NEF 3GPP components.

The other aspect of the threat surface are the external facing interfaces such as Peering points and roaming interfaces which the Mobile Network Operators today use for interconnection between the operators to allow their subscribers to roam between them. Roaming agreements are operator specific and there are restrictions on the roaming traffic to be terminated locally and the GTP tunnel is established from the roaming partners core in the existing networks - any malicious attack on the roaming partners NW would compromise the Host network.

Use of Third party networks to backhaul traffic / Lease backhaul from Third party networks causes the Gi / SGi interface to be integrated with some third-party networks which have been compromised and leaves the Mobile Network to be attacked.

There are also regulatory aspects on security which will also apply to the 5G architecture. For example, in Europe, the new General Data Protection Regulation (GDPR) will come into effect in May 2018. In order to comply with the GDPR, any company which collects, stores and processes personal data (i.e., relating to an identified or identifiable natural person) has a number of obligations. Failure to comply with the GDPR can incur hefty fines. Different components in the existing and the evolving 5G architecture would be interacting with personal data on multiple levels, only a secure, privacy and threat centric approach to 5G architecture can ensure the conformity to GDPR.

Mitigating Threats in 5G and Evolving Architectures using Cisco Portfolio

The evolution of the packet core in the mobile network operator architecture brings in need for securing the distributed UP components, security for the network slice, securing distributed data center, DDoS security and the API security as the packet core functions can now be reached by external applications via an API, using Hardened NFV infrastructure, requirement for enhanced visibility on the traffic flows within the packet core traffic flow, enhanced DDoS protection to prevent DDoS attacks on packet core, better segmentation and NGFW capabilities for Central DC / packet core and enhanced user access control as shown below.

The control plane and user plane separation further aids the distributed deployment model. For the Mobile Network packet core nodes, virtualization brings in the flexibility to deploy it in a distributed fashion in multiple datacenters and the functions can now be divided into centralized and distributed functions, depending on where its best placed for better efficient use of network resources. Security for virtualized function and NFV largely depends on isolation between the virtualized components brought in by segmentation and lifecycle management of the virtual functions deployed on the NFVI, which is brought in by security methods such as secure boot, trust anchor etc.

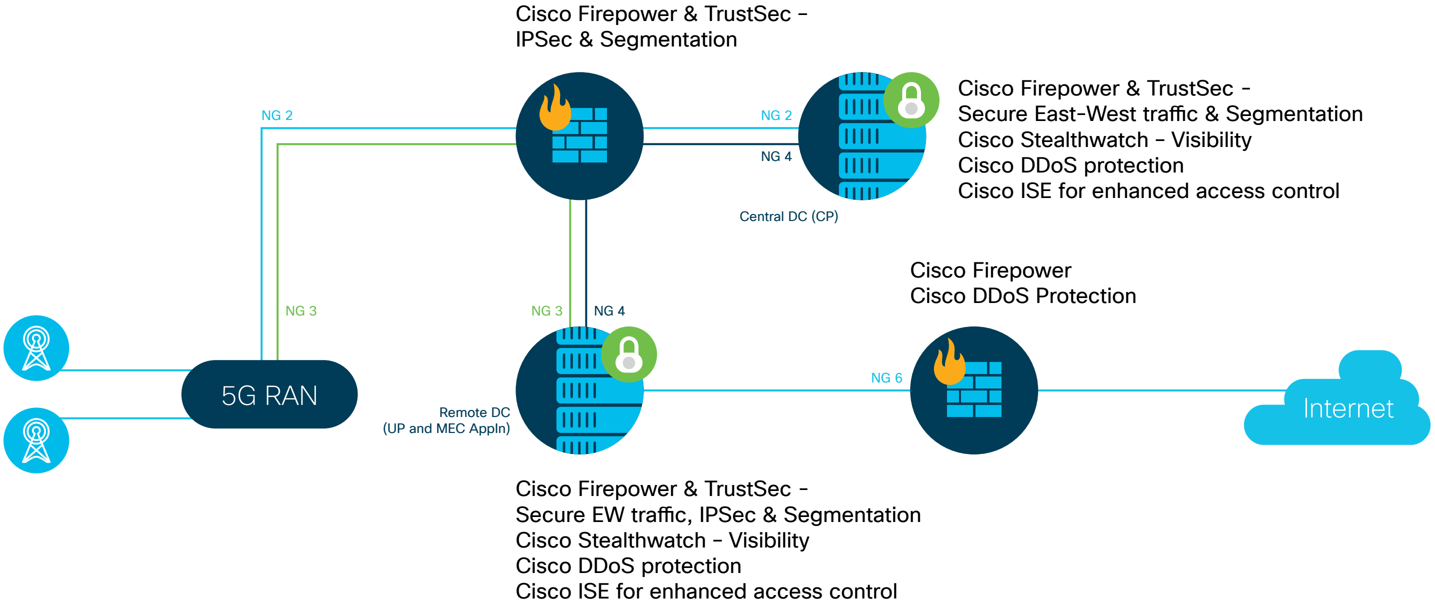

Cisco Firepower & TrustSec –
IPSec & Segmentation
NG 2
NG 2
NG 4
Cisco Firepower & TrustSec –
Secure East-West traffic & Segmentation
Cisco Stealthwatch – Visibility
Cisco DDoS protection
Cisco ISE for enhanced access control
Central DC (CP)
Cisco Firepower
Cisco DDoS Protection
NG 3
NG 3
NG 4
5G RAN
Remote DC
(UP and MEC Appln)
NG 6
Internet
Cisco Firepower & TrustSec –
Secure EW traffic, IPSec & Segmentation
Cisco Stealthwatch – Visibility
Cisco DDoS protection
Cisco ISE for enhanced access control

Figure 7. Cisco Security Portfolio Applied to the 5G Architecture

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 to prevent the vulnerabilities to spread to other components within the slice and between the slices in case of any malicious attacks.

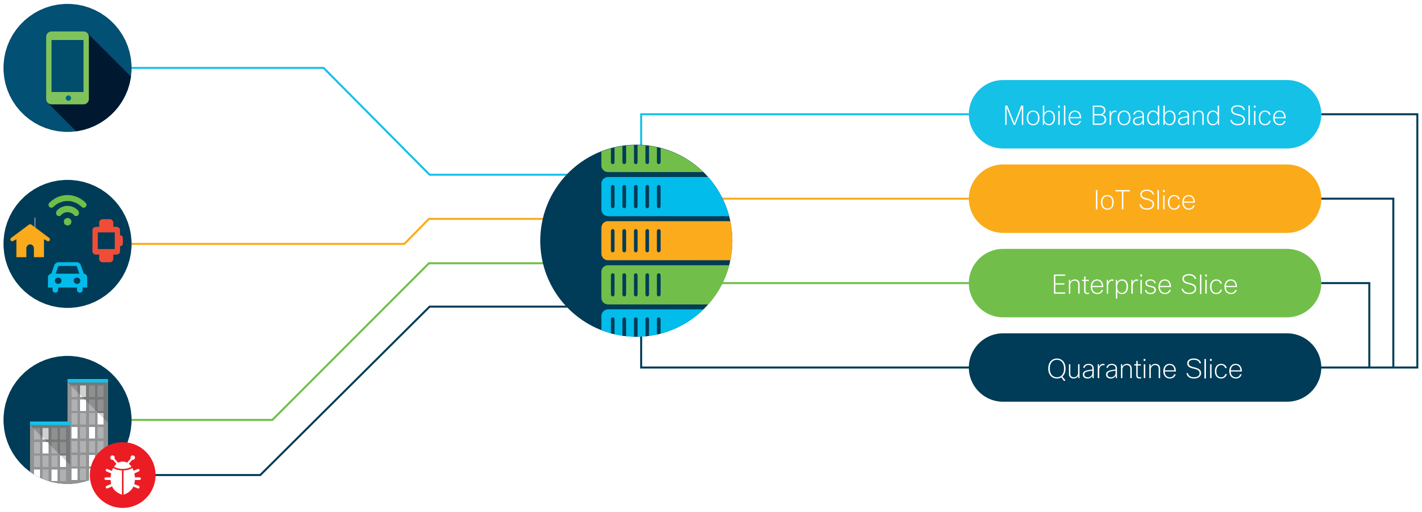


Figure 8. Threat Mitigation through Network Slicing

The network slices should also allow a quarantine slice which would enable the intra and inter slice security as shown below. The above implementation of the quarantine slice is made possible in transport by segment routing and in the data center by Cisco segmentation technology. Cisco's architecture joins these two segmentation methods and delivers an end to end segmented network delivering visibility and agility in threat detection and management. Together, segment routing and Trustsec deliver Software Defined Segmentation. Software defined segmentation makes it possible to enforce the access policies for users, applications and devices, which can apply to IoT devices, M2M based devices and enterprise network devices as well. In the Data Center part of the 5G architecture, software defined segmentation leverages segment routing and when TrustSec is used, the security group tags can be defined and managed by Cisco's Identity Services Engine which can also share TrustSec group information with other group-based policy schemes allowing segmentation of the traffic and restricting communication between defined network interfaces. This security approach would shift the network security away from depending on long lists of IP address to a flexible, automated model which is better managed and more effective against new and expanding threat vectors.

Evolving backhaul and RAN architectures brings in the need for enhanced device security which includes Endpoint Protection and advanced Malware Protection, DNS based protection and scrubbing, secure control and management of device for IoT, DDoS security and distributed SecGW functionality to cater for interface protection between the eNB and gNB and between gNB's, blocking devices from reaching out to malicious servers and RAN vendor controlled techniques for anti- jamming for securing the air interface vulnerabilities.

For the protection of the MEC servers and distributed datacenters, Cisco recommends backhaul and distributed DC threat mitigation using enhanced segmentation by using software defined segmentation and user access control, DDoS protection for remote DC where MEC server and vBBU would be deployed and enhanced visibility on the network flows within the remote DC.

Securing datacenter and cloud components are becoming critical as the mobile network components are being virtualized and can be deployed on an NFVI or as a microservice component in the cloud. For securing the Datacenter, all the components within the Datacenter should be secured apart from the securing the perimeter. The related components are the NFVI infrastructure security which relates to hardening the NFVI hardware, securing E-W security, VNF & container security such as isolation between the VNF's, detecting malicious behavior of the virtual functions, securing the third-party application and API, securing & segmenting the network interfaces, roaming and peering interfaces, and then securing the user access and the orchestration layer.

SDN/NFV Security

Constraints are delivered in many different ways across the solution. When looking at the data center or cloud used to deliver virtual functions to form a systemic approach to deliver services, segmentation is required to insure security separation, visibility and controls. The automation of these security controls, the infrastructure changes they protect and other key network operations and security operations tasks is a key feature of the solution provided. The 8 step process used to secure SDN/NFV is described in the graphic below.

The security architecture leverages a foundation of NFVi to provide a layer in the architecture which all other vertical applications (be they EPC or other) plug in to. Therefore, the eight step process described in the graphic below to secure the NFVi and the applications that run on top of it and are orchestrated for it, provides the systemic approach to security in this context (for slicing). Specific features that are part of each network slice will have their own domain specific functions and security concerns, but the overall approach to controlling the "impact domain" for network slicing is described in the graphic below.

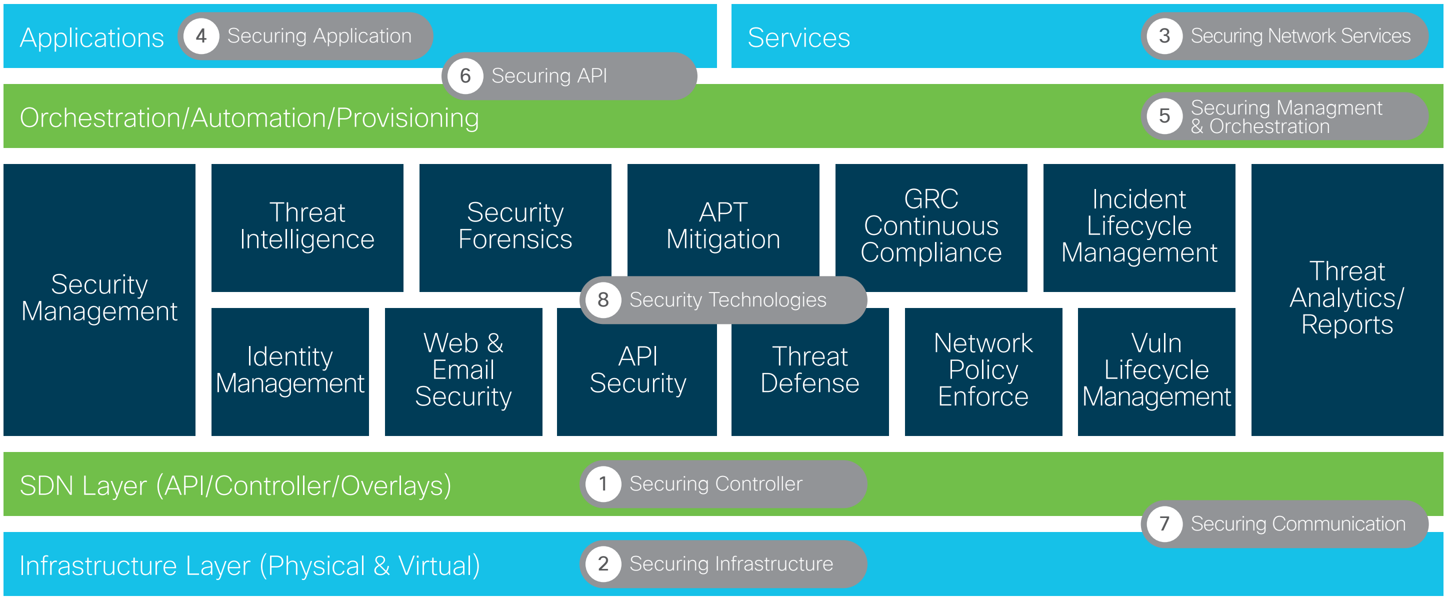
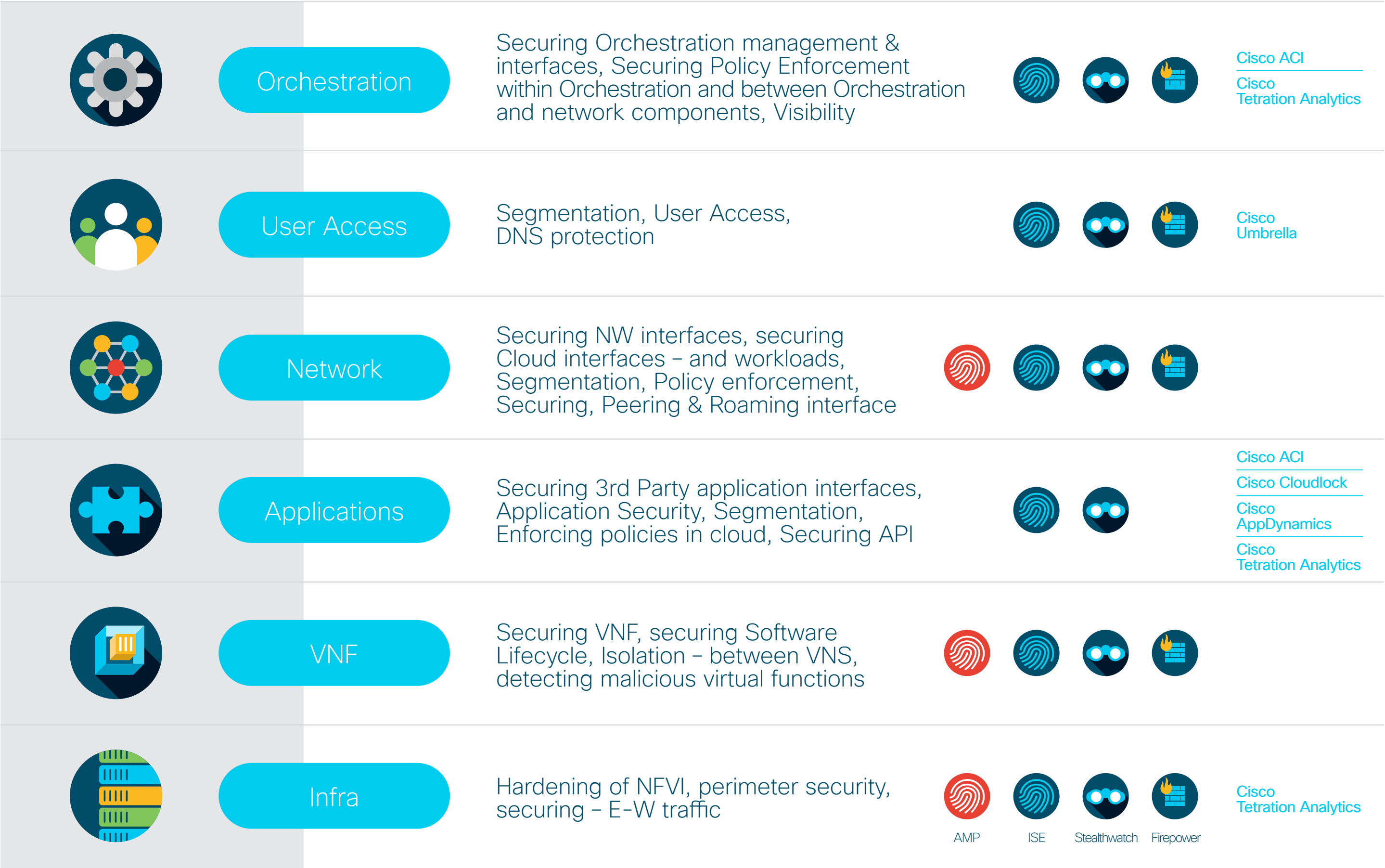

Applications
4 Securing Application
Services
3 Securing Network Services
6 Securing API
Orchestration/Automation/Provisioning
5 Securing Managment & Orchestration
Security Management
Threat Intelligence
Security Forensics
APT Mitigation
GRC Continuous Compliance
Incident Lifecycle Management
Threat Analytics/ Reports
8 Security Technologies
Identity Management
Web & Email Security
API Security
Threat Defense
Network Policy Enforce
Vuln Lifecycle Management
SDN Layer (API/Controller/Overlays)
1 Securing Controller
7 Securing Communication
Infrastructure Layer (Physical & Virtual)
2 Securing Infrastructure

Figure 9.

In today's networks we deliver application outcomes and those applications need to securely ask the "network" for a specific set of behaviors. The process of securing the application's requests to the network, treating the network programmatically to deliver an outcome, is detailed in the graphic above. There are 8 layers of the threat vectors associated with securely delivering a service in a network evolved to include SDN, NFV and virtualization. This is not meant to be a comprehensive list of threats, but more a group of related categories that represent them.

Cisco provides multi layered security components to secure the end to end architecture, including the securing the Orchestration components for data center and telco cloud as illustrated in the Figure 10.



		AMP	ISE	Stealthwatch	Firepower	
Orchestration	Securing Orchestration management & interfaces, Securing Policy Enforcement within Orchestration and between Orchestration and network components, Visibility		●	●	●	Cisco ACI Cisco Tetration Analytics
User Access	Segmentation, User Access, DNS protection		●	●	●	Cisco Umbrella
Network	Securing NW interfaces, securing Cloud interfaces – and workloads, Segmentation, Policy enforcement, Securing, Peering & Roaming interface	●	●	●	●	
Applications	Securing 3rd Party application interfaces, Application Security, Segmentation, Enforcing policies in cloud, Securing API		●	●		Cisco ACI Cisco Cloudlock Cisco AppDynamics Cisco Tetration Analytics
VNF	Securing VNF, securing Software Lifecycle, Isolation – between VNS, detecting malicious virtual functions	●	●	●	●	
Infra	Hardening of NFVI, perimeter security, securing – E-W traffic	●	●	●	●	Cisco Tetration Analytics

Figure 10.

Securing the orchestration layer becomes very important in the evolving architectures as the networks would be software based with a split between hardware and software. Therefore, cloud and virtualized telecommunication technology, SDN, industry 4.0 (Internet of Things), etc. will be taken into consideration for the 5G deployment which has impact on the network management and orchestration.

The graphic above shows the layered approach to delivering and operating the distributed data center infrastructure that 5G and its services will run on. At the bottom layer, visibility is provided into application white listing and traffic patterns via Cisco Tetration, giving the operator visibility into patterns to baseline normal and to deliver a zero trust or application white list approach for secure connectivity. Tetration also offers a network DVR of sorts which can play back what was happening in that infrastructure at any given time. On top of the Infrastructure layer is the NFV infrastructure providing orchestration and service delivery and placement. The Cisco NFV infrastructure supports multi-vendor orchestration, service placement and assurance and advanced networking to services in the data center and the cloud. The Infrastructure layer and the NFV Infrastructure layer together provide a horizontal foundation. All services insert themselves as a vertical pillar into the horizontal foundation. This is what makes it possible to place the security controls in the right place, at the right time, as close to the source of the threat as possible because in the orchestration function, Cisco is able to leverage a tight integration with the network to deliver optimal workload placement, including at distributed data center nodes.

The orchestration layer would manage legacy components as well as the 5G components of the architecture and also would interact or contain the security orchestration functions for the architecture as well, as shown in Figure 11.

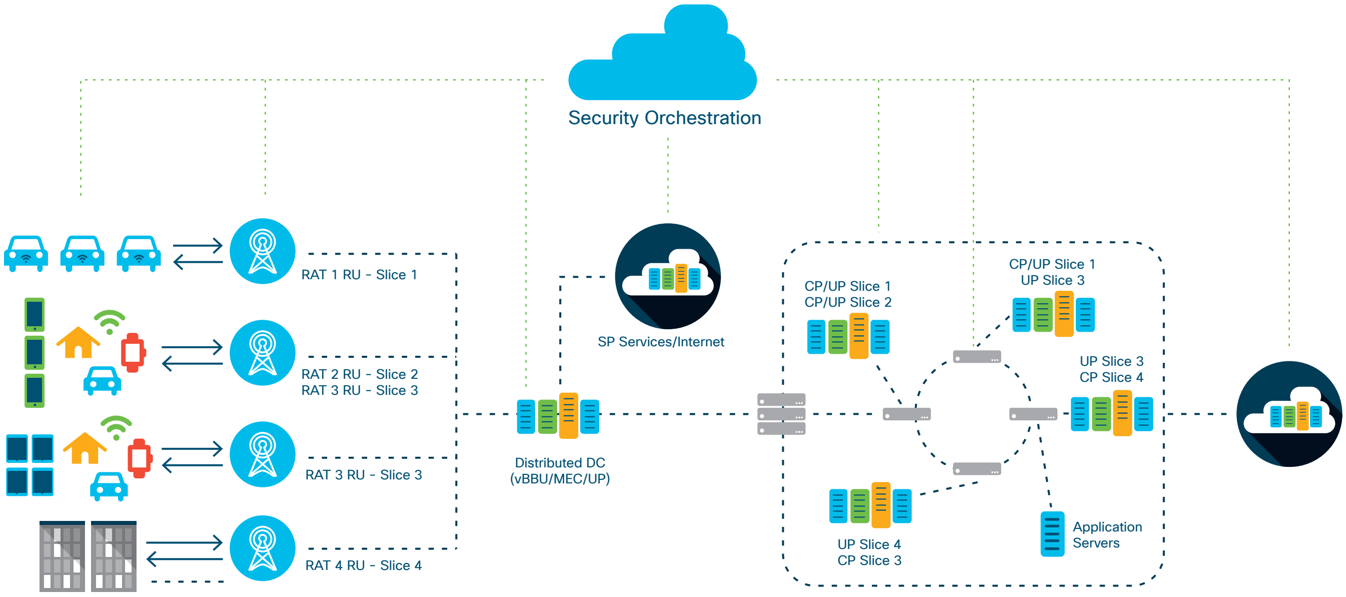


Figure 11.

This brings in the need for enhanced security for the orchestration layer such as having proper policies, segmentation, user access control and better visibility and care should be taken in architecture for orchestration for preventing a single point of failure. This can be achieved as shown below by allowing closed loop orchestration architecture where security is applied not only on the orchestration layer to allow the policies to be applied globally but also to individual components, which can be achieved by a variation of the mesh architecture as shown in Figure 12.

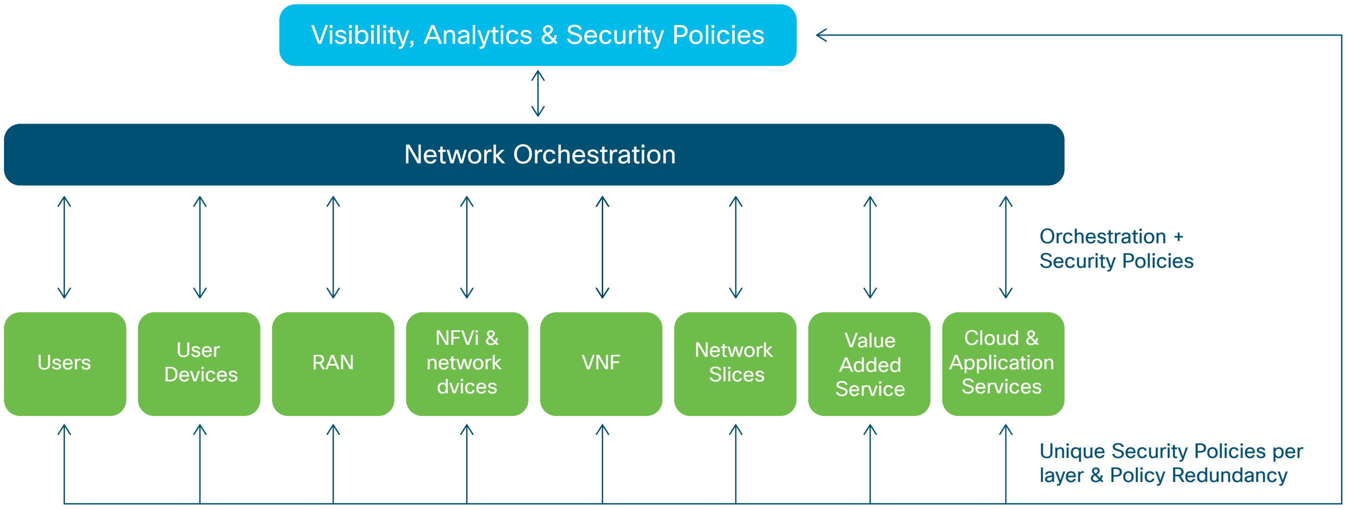


Figure 12.

Using the virtualization benefits of scalability and efficient use of resources a step further are the Telco cloud architectures which allow faster service deployment and quicker improvement in existing cloud based microservice components deployed by the mobile network operators as shown in Figure 13.

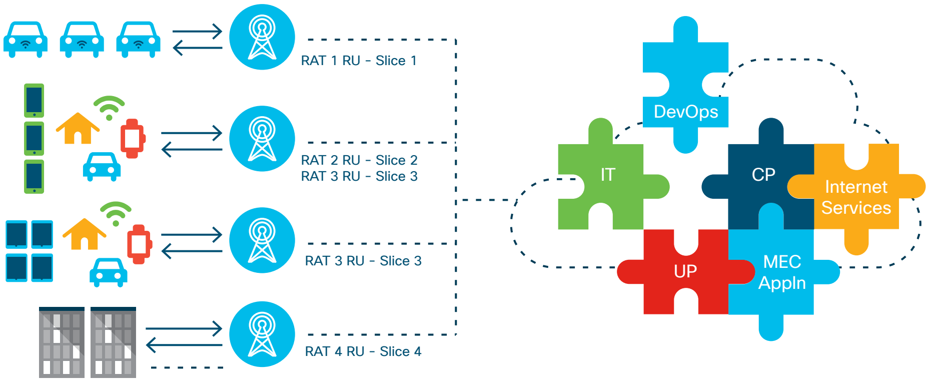


Figure 13.

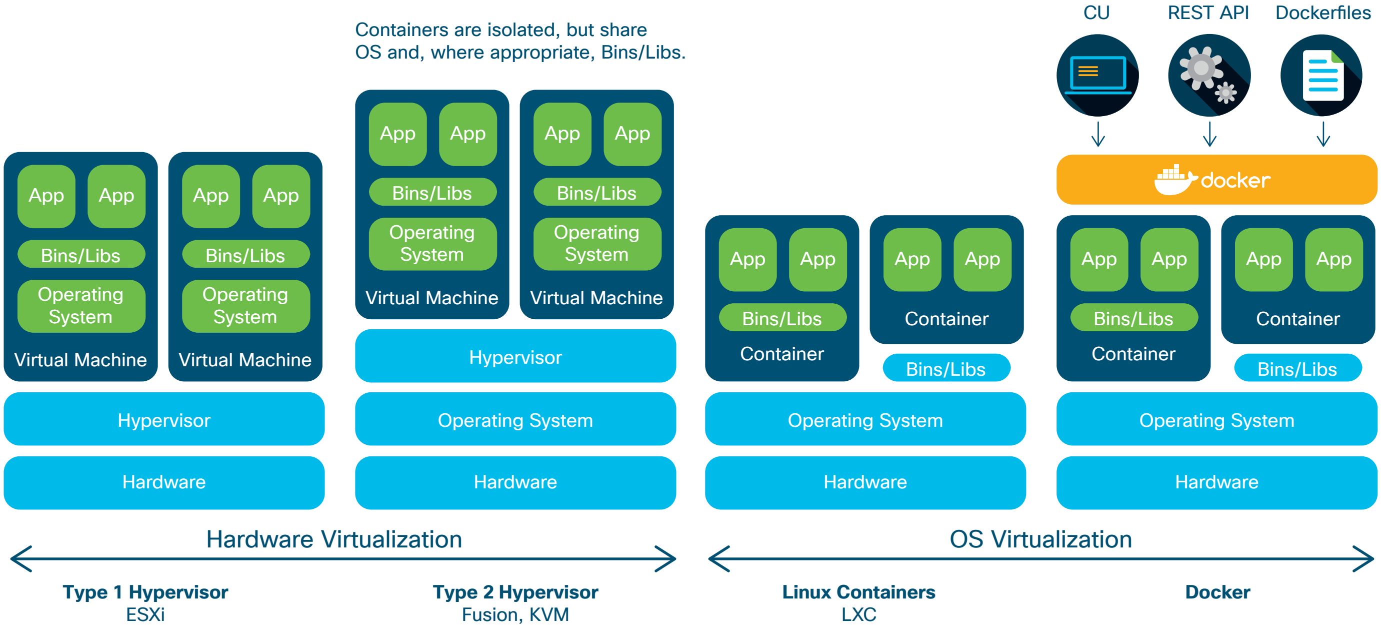
Cloud deployment of the 5G network components brings in security requirements for the microservice components in the cloud infrastructure which are being updated regularly by the DevOps team and the security policies being applied to the services which are being scaled in and scaled out continually depending on the capacity requirements. Cisco's Stealthwatch Cloud can allow enhanced visibility of the flows between 5G components and mitigate threats from malicious traffic within the Telco cloud architecture.

Because 5G will run as a set of virtualized services, orchestrated micro services and containers, we must then consider the security implications of the DevOps and more specifically the SecDevOps model. Cisco utilizes the CSDL or Cisco Secure Development Lifecycle to ensure that development, testing and delivery are all executed timely and meet security gates to pass key development milestones. CSDL insures that key trust boundaries at the network and applications that run on top of them are fully protected by documented processes and gates.

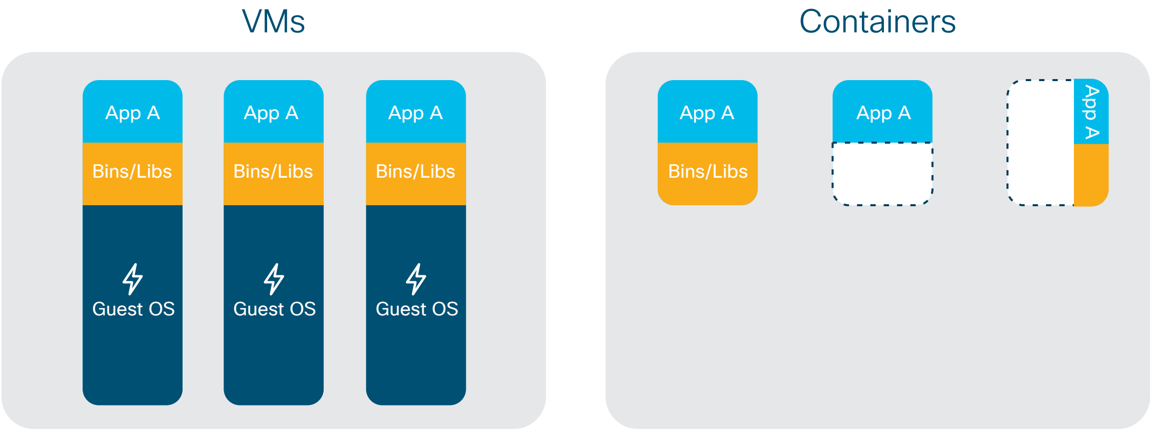
Micro services functions and characteristics are detailed below.

- Split the application into a set of small, loosely interconnected services (micro services)
- Each service implements a separate function of the application such as order management, reporting, payments, etc.
- Communication between micro-services through APIs such as REST
- Allows for CI/CD as separate development teams can work independently due to loose coupling and push out rapid code releases.
- Often these microservices are packaged in linux or docker containers.
- Containers should be stateless – data stored in separate layer of the architecture (such as Cassandra)

Securing traffic between virtual machines is more advanced in capabilities today than security containers. Techniques for visibility and controls at the micro service level for containers are here but require operational BCPs (best common practices). The graphic below highlights the difference between the use of vm's vs. containers.



Containers and micro services are the preferred path for virtualization of workloads going forward because of how much more operationally efficient and "lighter" that containers are versus virtual machines. The graphic below shows why.



VMS
Every app, every copy of an app and every slight modification of the app requires a new virtual server

Original App
(No OS to take up space, resources, or require restart)

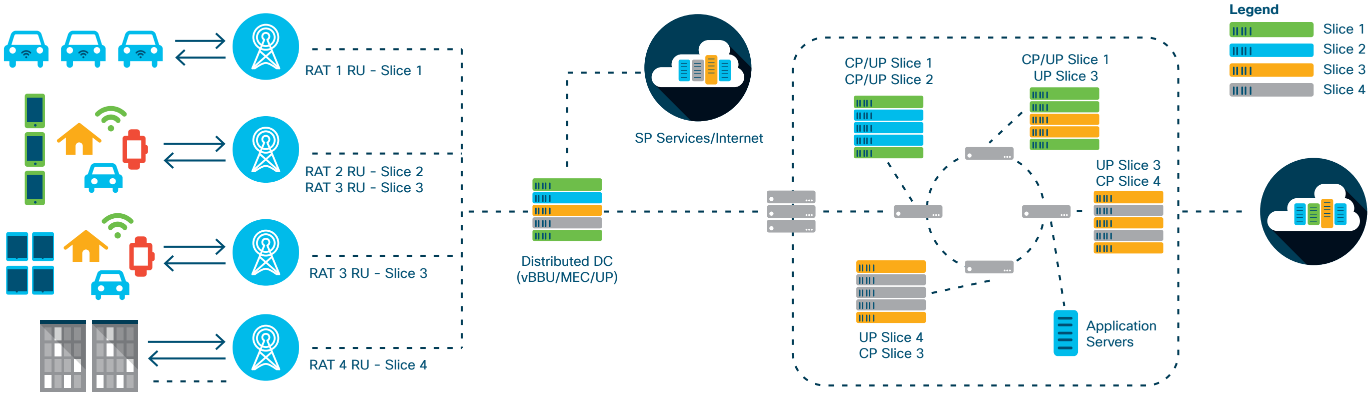
Copy of App
No OS. Can share bins/libs

Modified App
Union file system allows us to only save the diffs between container A and container A

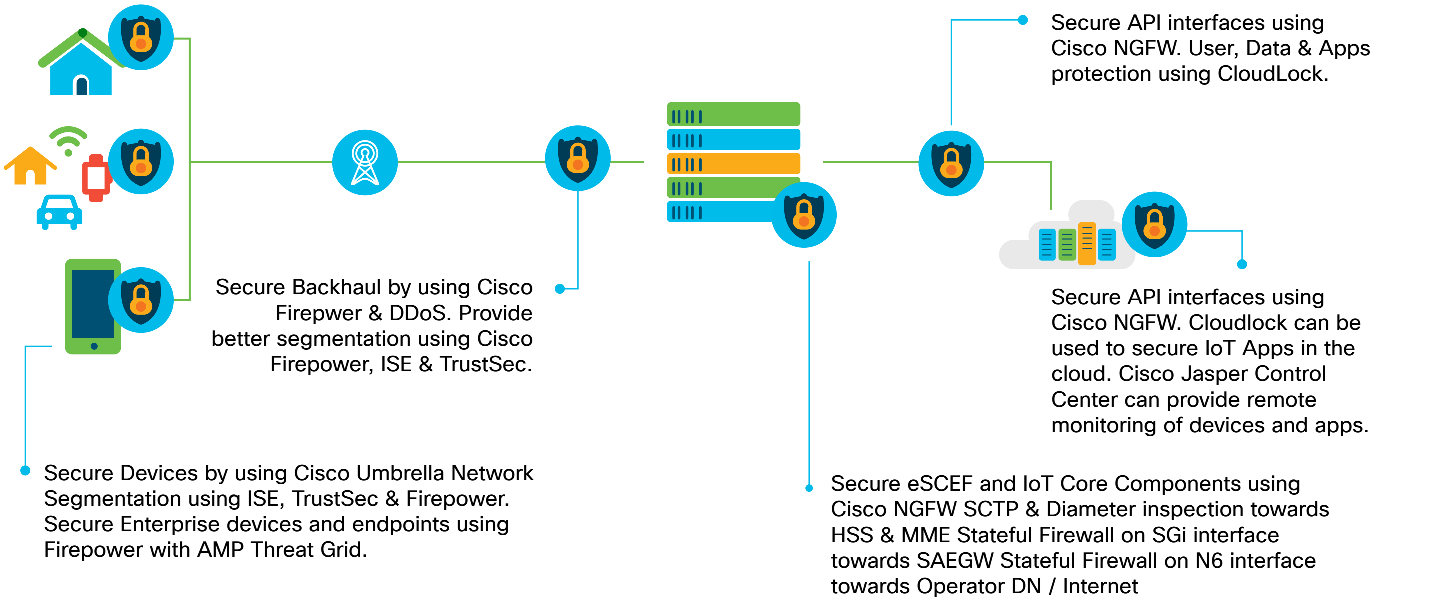
When thinking about how to secure services, the perimeter and trust boundaries all must be secured, but there is a complement of security approaches and controls for security the DevSecOps element as well. These include:

- App Security integrated with CI/CD
  - Static code analysis (Veracode is an example)
  - Dynamic analysis (Web App Inspector - Fortify WebInspect)
- Automation Infrastructure
  - Lock down the automation and orchestration tools
  - Lock down keys and roles (Ansible with Vault)
- Infrastructure Security
  - Configuration management and hardening (Terraform, Ansible, Puppet)
  - Docker security verification during the build process

Cisco provides an end to end Security architecture for securing Mobile Network Operators evolving network and the above figure shows the different components which can be deployed to secure the network.



	Device Threats	Air Interface Threats	RAN Threats	Backhaul Threats	5G Packet Core & OAM Threats	SGI/N6 & External Roaming Threats
Stealthwatch	●	●	●	●	●	●
Umbrella	●	●	●	●	●	●
Tetration				●	●	●
Firepower			●	●	●	●
ISE + TrustSec	●	●	●	●	●	●
AMP	●			●	●	



Secure API interfaces using Cisco NGFW. User, Data & Apps protection using CloudLock.

Secure Backhaul by using Cisco Firepwer & DDoS. Provide better segmentation using Cisco Firepower, ISE & TrustSec.

Secure API interfaces using Cisco NGFW. Cloudlock can be used to secure IoT Apps in the cloud. Cisco Jasper Control Center can provide remote monitoring of devices and apps.

Secure Devices by using Cisco Umbrella Network Segmentation using ISE, TrustSec & Firepower. Secure Enterprise devices and endpoints using Firepower with AMP Threat Grid.

Secure eSCEF and IoT Core Components using Cisco NGFW SCTP & Diameter inspection towards HSS & MME Stateful Firewall on SGi interface towards SAEGW Stateful Firewall on N6 interface towards Operator DN / Internet

Figure 15. Mitigation of IoT Threats

Threat and Dark Threat Analysis

At this stage of this white paper, you now recognize how important visibility of traffic is to the proper operation of security of the 5G network. What would you do if that visibility is taken away? How could you then determine that traffic is malicious or contains malware? The answer is Cisco ETA or encrypted traffic analytics. From https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/nb-09-encrytd-traf-anlytcs-wp-cte-en.pdf, let's dig into Encrypted Traffic Analytics.

The rapid rise in encrypted traffic is changing the threat landscape. As more businesses become digital, a significant number of services and applications are using encryption as the primary method of securing information. More specifically, encrypted traffic has increased by more than 90 percent year over year, with more than 40 percent of websites encrypting traffic in 2016 versus 21 percent in 2015. Gartner predicts that by 2019, 80 percent of web traffic will be encrypted.

Encryption technology has enabled much greater privacy and security for enterprises that use the Internet to communicate and transact business online. That same encryption is making the network operator's ability to see the traffic and determine if it's malicious much more difficult. Mobile, cloud and web applications rely on well-implemented encryption mechanisms, using keys and certificates to ensure security and trust. However, businesses are not the only ones to benefit from encryption. Threat actors have leveraged these same benefits to evade detection and to secure their malicious activities.

Visibility across the network is getting increasingly difficult and our traditional means of detection cannot assume that data is available for inspection. We need to be able to simultaneously assess how much of our digital business is protected and unprotected by encryption while also assessing what traffic is malicious and what is benign. Gartner believes that half of malware campaigns in 2019 will use some type of encryption to conceal malware delivery, command and control activity, or data exfiltration.

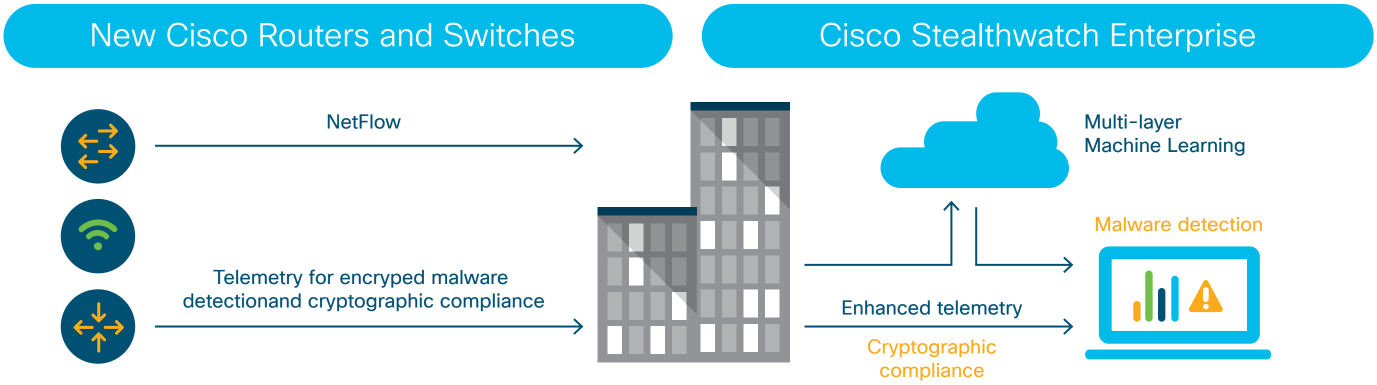
The majority of organizations today do not have a solution to detect malicious content in encrypted traffic. They lack the security tools and resources to implement a solution that can be deployed throughout their network infrastructure without slowing down the network.

Traditional threat inspection with bulk decryption, analysis and re-encryption is not always practical or feasible, for performance and resource reasons. In many cases, however, advanced analytic techniques can be used to identify malicious flows for further inspection using decryption techniques.

On any given day, no one knows how much of their digital business is in the clear versus encrypted. If traffic is encrypted, the encryption is typically done to meet compliance requirements that mandate specific security policies.

Traditional flow monitoring provides a high-level view of network communications by reporting the addresses, ports and byte and packet counts of a flow. In addition, intra-flow metadata, or information about events that occur inside of a flow, can be collected, stored and analyzed within a flow monitoring framework. This data is especially valuable when traffic is encrypted, because deep-packet inspection is no longer viable. This intra-flow metadata, called Encrypted Traffic Analytics, is derived by using new types of data elements or telemetry that are independent of protocol details, such as the lengths and arrival times of messages within a flow. These data elements have the attractive property of applying equally well to both encrypted and unencrypted flows.

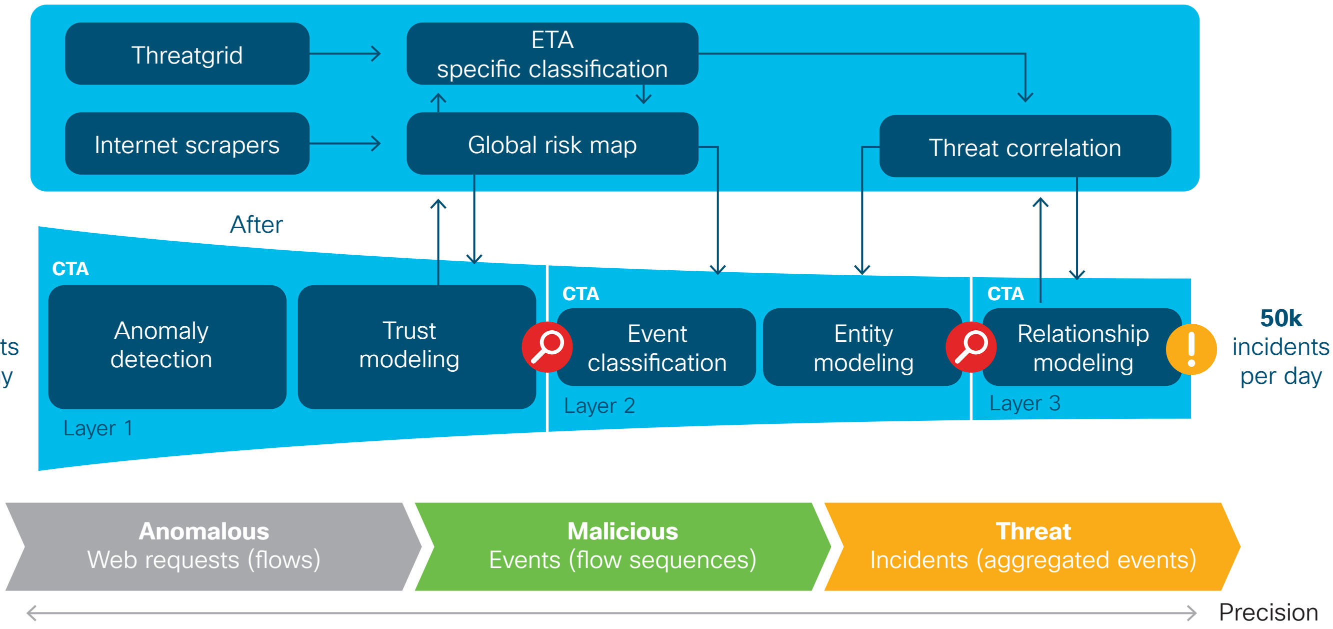
Using these data elements or intraflow telemetry to identify malware communication in encrypted traffic means Encrypted Traffic Analytics can maintain the integrity of the encrypted flow without the need for bulk decryption.

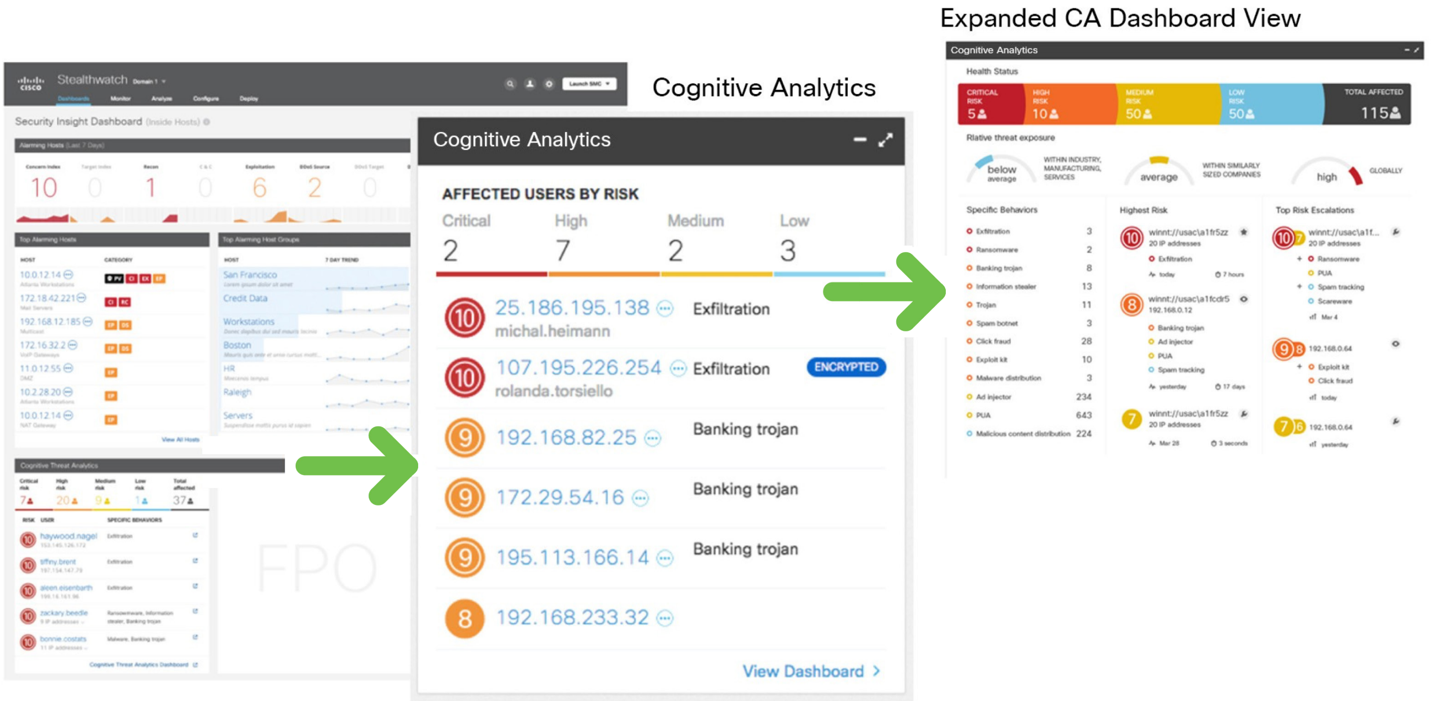

New Cisco Routers and Switches
Cisco Stealthwatch Enterprise
NetFlow
Telemetry for encryped malware
detectionand cryptographic compliance
Multi-layer
Machine Learning
Malware detection
Enhanced telemetry
Cryptographic
compliance

Leveraged network	Faster Investigation	Higher precision	Stronger protection
Enhanced telemetry from Cisco’s newest switches and routers	Advanced analytics and machine learning	Global-to-local knowledge correlation	Continuous Enterprise-wide compliance

Interestingly, the same visibility platform referenced earlier in this white paper, Stealthwatch, is the same that is the core of the ETA functionality. Continuity and integration across the tools that provide the security fabric for 5G networks is a key tenet of the Cisco Security Architecture for 5G.

The following two diagrams highlight key functionality, the role of machine learning and how the various elements of the ETA process work together.





Security Insight Dashboard (Inside Hosts)

Alarming Hosts (Last 7 Days)

Concern Index	Target Index	Recon	C & C	Exploitation	DDoS Source	DDoS Target
10	0	1	0	6	2	0

Top Alarming Hosts

HOST	CATEGORY
10.0.12.14	PV CI EX EP
172.18.42.221	CI RC
192.168.12.185	EP DS
172.16.32.2	EP DS
11.0.12.55	EP
10.2.28.20	EP
10.0.12.14	EP

View All Hosts

Top Alarming Host Groups

HOST	7 DAY TREND
San Francisco	
Credit Data	
Workstations	
Boston	
HR	
Raleigh	
Servers	

Cognitive Threat Analytics

Critical risk	High risk	Medium risk	Low risk	Total affected
7	20	9	1	37

RISK	USER	SPECIFIC BEHAVIORS
10	haywood.nagel	Exfiltration
10	tiffny.brent	Exfiltration
10	aleen.eisenbarth	Exfiltration
10	zackary.beedle	Ransomware, Information stealer, Banking trojan
10	bonnie.costats	Malware, Banking trojan

Cognitive Threat Analytics Dashboard

Cognitive Analytics

AFFECTED USERS BY RISK

Critical	High	Medium	Low
2	7	2	3

10	25.186.195.138 michal.heimann	Exfiltration	
10	107.195.226.254 rolanda.torsiello	Exfiltration	ENCRYPTED
9	192.168.82.25	Banking trojan	
9	172.29.54.16	Banking trojan	
9	195.113.166.14	Banking trojan	
8	192.168.233.32		

View Dashboard >

Health Status

CRITICAL RISK	HIGH RISK	MEDIUM RISK	LOW RISK	TOTAL AFFECTED
5	10	50	50	115

Rlative threat exposure

below average WITHIN INDUSTRY, MANUFACTURING, SERVICES
average WITHIN SIMILARLY SIZED COMPANIES
high GLOBALLY

Specific Behaviors

Exfiltration	3
Ransomware	2
Banking trojan	8
Information stealer	13
Trojan	11
Spam botnet	3
Click fraud	28
Exploit kit	10
Malware distribution	3
Ad injector	234
PUA	643
Malicious content distribution	224

Highest Risk

10 winnt://usac\a1fr5zz
20 IP addresses
Exfiltration
today 7 hours

8 winnt://usac\a1fcdr5
192.168.0.12
Banking trojan
Ad injector
PUA
Spam tracking
yesterday 17 days

7 winnt://usac\a1fr5zz
20 IP addresses
Mar 28 3 seconds

Top Risk Escalations

10 7 winnt://usac\a1f...
20 IP addresses
+ Ransomware
PUA
+ Spam tracking
Scareware
Mar 4

9 8 192.168.0.64
+ Exploit kit
Click fraud
today

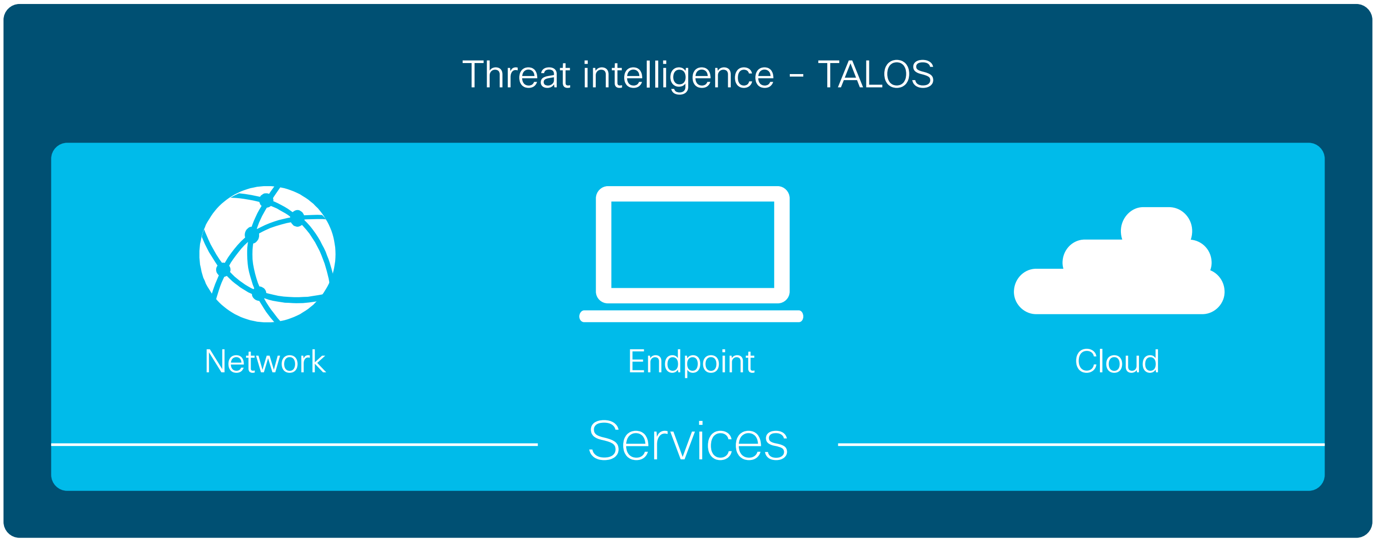
7 6 192.168.0.64
yesterday

ETA is a great example of why security and the network that it is protecting have to be integrated at every step. Application, cloud, data center, network and endpoint are all treated as a secured integrated system. The integration of the unit parts, from a security perspective is described earlier in this white paper.

Unmatched Threat Analytics and Intelligence

The Talos team protects your organization’s people, data, and infrastructure. Our researchers, data scientists, and engineers collect information about existing and developing threats. We then deliver protection against attacks and malware. Talos underpins the entire Cisco security ecosystem. Cisco research and threat services not only empowers our platforms, but also is a key aspect of integration with an operator, making that operator’s threat systems that much more effective, both in incident response and other key Security Operations Center functions.

Talos provides end to end threat intelligence and research:



Talos research is comprehensive and keeps the operator and the operator’s customers ahead.

Hundreds of Thousands Customers

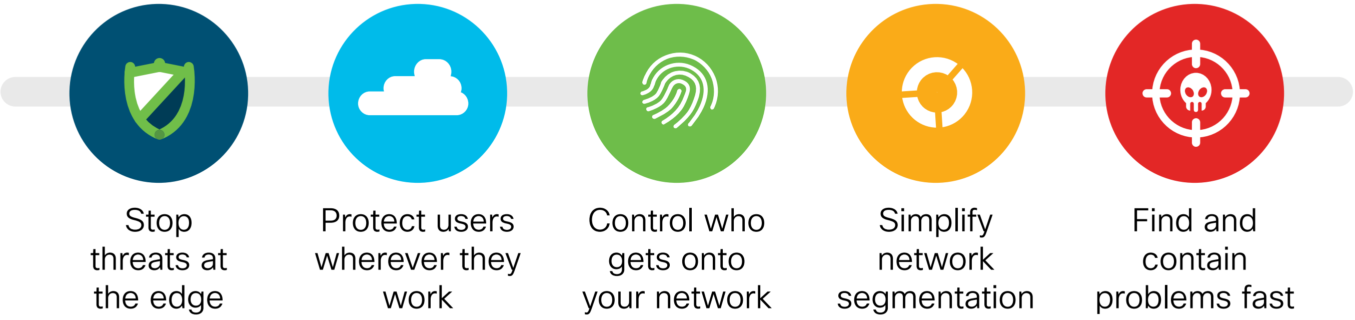
Tens of Millions Users

19.7B Threats Blocked Per Day

300 Threat Researchers

Hundreds of Threat Analytic Engines

In conclusion, security for 5G is an integrated approach with visibility and control from end to end. The focus areas covered today and summarized in the graphic below:



Stop threats at the edge

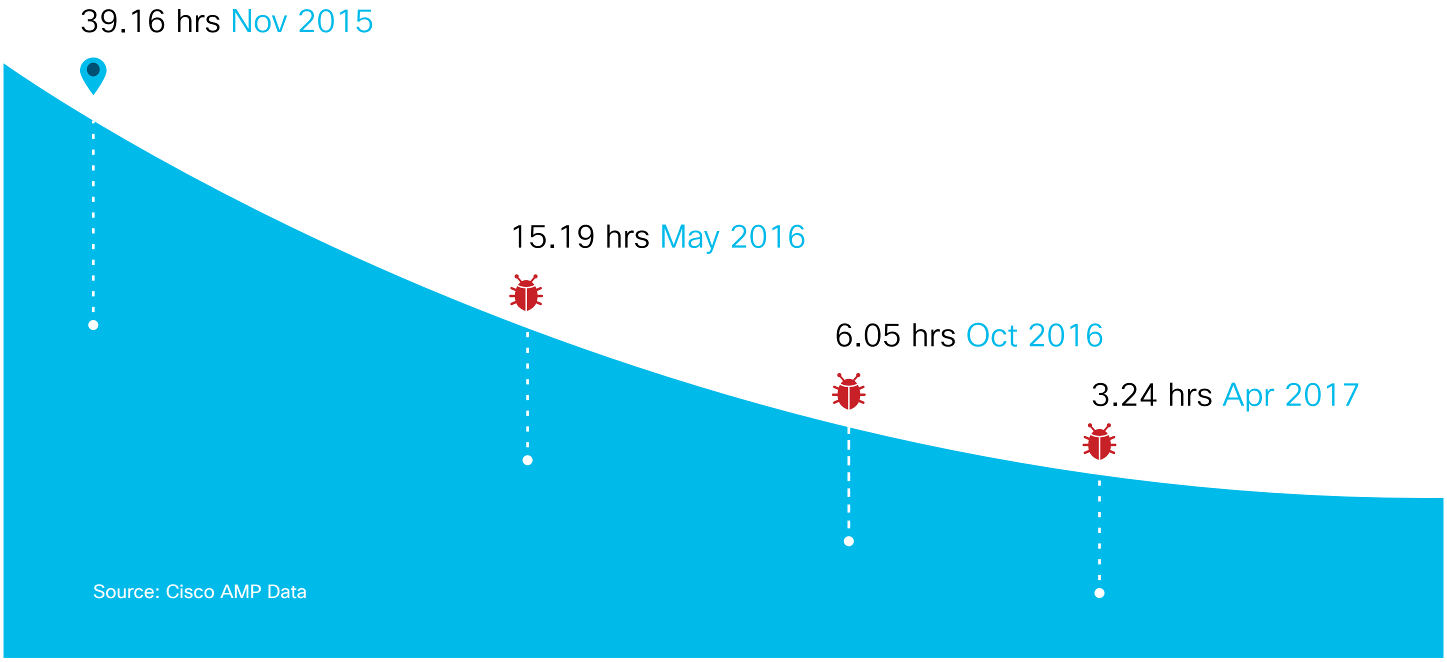
Protect users wherever they work

Control who gets onto your network

Simplify network segmentation

Find and contain problems fast

The mean time to detection using Cisco security technology continues to go down enabling the operator better agility for service deployment with confidence.



Source: Cisco AMP Data