MAVENIR

Open architecture and supply chain diversity: securing telecoms into the future

White paper

JANUARY 2023

Contents

1. Exec Summary

1 Taking Wireless Telecoms into the Cloud

The technology that fuels connectivity powers some of the greatest innovations known to mankind – yet the supporting infrastructure behind it is failing to keep pace. Currently, wireless computing is largely carried out at the tower site – processing data, encrypting it, receiving, or transmitting it through radio signals. Essentially, that means wireless infrastructure is still in its IBM era of mainframe computing, where one whole floor of the office building is dedicated to the company server.

The majority of this computing power could be run remotely in a cloud environment, unlocking significant benefits when it comes to flexibility, cost savings and network interoperability. Yet, with wireless infrastructure dominated by legacy providers interested in maintaining the status quo, the sector is falling behind in the transition to a cloud-based infrastructure that is rapidly being adopted by other industries. The movement towards open networks in telecoms is disruptive – and like all disruptors invites criticism, particularly in relation to security.

However, less well explored so far are the major security advantages of open, interoperable networks, including significantly reducing exposure to geopolitical risk, creating safer physical environments for data processing, expanding the supply chain to reduce vulnerabilities and vendor lock in, as well as enhanced resilience to cyber-attack.

These factors are explored in this paper and inform Mavenir's five key security principles designed to drive the transition to open systems with security front of mind.

2 The critical security benefits of cloud-based systems

Hedging geopolitical risk

The security of critical telco infrastructure has been discussed at great length over the last few years, prompted in large part by concerns over whether the Chinese government may have access to sensitive data via the largely state-owned technology giant Huawei.

Ensuring world-wide data flows remain secure is unanimously recognized as a strategic priority for governments, companies and the telco industry itself. However, delivering on this priority is significantly hampered with the continued development and manufacturing of equipment in China.

Annual reports from European telco incumbents demonstrate a significant investment within China in recent years, indicating risk at the earliest stages of the supply chain which essentially negate the impacts of 'Rip & Replace'.

Less well understood are the geopolitical risks beyond China. The invasion of Ukraine by the Russian Federation has unleashed geopolitical and economic chaos far beyond the theatre of war. Both Nokia and Ericsson announced that they are winding down operations in Russia by the end of 2022, but their Nordic homebases are still firmly in the crosshairs of their active neighbor.

The hot physical conflict in Eastern Europe is accompanied by a vast, spiraling cyber conflict across the globe. Geopolitical tensions are stoking hacking operations, malicious attacks to banks, networks, airlines and hospitals – all growing at an exponential rate. Compounding that is the risk of physical attacks on critical hardware unsecured across tower sites.

A diverse, competitive and virtualized supply chain is a vital hedge to geopolitical risk. Not locked in on hardware, flexible enough to change suppliers or elements of the IT operation, operators can shop around geographically.

Questions about Chinese hardware are moot in a virtualized context – there is no complete network 'Rip and Replace' for open, interoperable networks running on software agnostic hardware. State aggression is equally less of a threat to critical infrastructure when the supply chain is less concentrated in hardware offerings.

Building resilience to malicious software attacks

The case of whether to allow Huawei to supply critical hardware becomes moot if most of the telco network is moved into a software-based cloud environment. However, that does not negate the risk of attacks by malicious actors. In fact, the number of known malware attacks rose 11 percent in the first half of 2022, mostly targeted at financial institutions, but not exclusively. Poor governance, state sanctioned or sponsored hacking operations, political instability against a backdrop of uncertainty – the perfect conditions for hacking to grow are all in play.

Criticisms of the open future have included an expanded surface attack area with more entry points for malicious actors. In fact, each part of the virtualized network is protected through firewalls, meaning that attack on one part of the network can be isolated from the rest, resulting in less critical breaches if they are successful. It is worth noting that the number of points of entry and the surface area of attack is essentially the same as in a hardware environment in real protected deployments.

The key differentiator with open systems is that a Zero Trust Architecture (ZTA) rooted in the principle of "never trust, always verify," is applied at every stage of the design and implementation process. When properly implemented, a ZTA enhances enterprise cybersecurity over traditional network perimeter-focussed security often used in hardware-based and closed solutions by reducing resource exposure to attackers and minimizing, or preventing, lateral movement within an enterprise should a host asset be compromised.

That is because the Zero Trust model does not assume trustworthiness, unlike the perimeter security model, which accepts securing the network edge as sufficient and assumes actors inside the operator's network are inherently trustworthy.

Furthermore, open architecture also addresses a key systemic gap found in some hardware-based and closed solutions – the lack of standard interfaces between components. Vendors who rely on proprietary Application Programming Interfaces (APIs) provided by partners suffer several issues which challenge system security. This includes relying on a single controlling vendor who can change the API , resulting in efforts to upgrade products, product compatibility issues, the need for custom wrappers to interface products to custom APIs, as well as a lack of independent test capabilities.

Open interfaces, with standardized, vendor-independent APIs address these issues, including enabling third party test capabilities. These and related concepts are described in the US National Institute of Standards and Technology (NIST) SP 800-207 Zero Trust Architecture (nist.gov). By leveraging network segmentation, preventing lateral movement, providing 'Layer 7' threat prevention, and simplifying granular user-access control, open architecture allows for much greater visibility through one system-wide interface.

Improving physical security

Now that lockdowns have largely abated, so too has the intensity of the conspiracy theories that fueled the attacks on critical telco infrastructure that partially characterized the early pandemic. The burning of telco towers is an extreme example of the physical risk to computing structures, but a serious and costly one.

The physical presence of so much data in relatively unsecured environments represents a major, under-researched and largely unaddressed threat. It is theoretically possible to physically siphon data from telco tower computing equipment. Plugging into ports would allow malicious actors to access sensitive data around major diplomatic meetings, for instance, or to access high value commercial information. There is evidence that backdoors were built into legacy tower infrastructure in Russia to allow the government to intercept communications.

Relocating racks off cell sites to data centers vastly improves their physical security. Computing offsite has major physical benefits as server farms that power the cloud are distributed and secured at unidentified locations. This makes them significantly less vulnerable to acts of vandalism and to physical siphoning of data via cables.

With Open RAN architectures, for instance, information encrypted at source is not decrypted at the radio site, meaning data is protected more comprehensively as it travels inside the carrier network.

Building diverse supply chains

Supply chain diversification is fundamental to resilience. In April 2022, following two years of pandemic related supply chain disruption, The International Monetary Fund called on businesses to diversify.

And yet, telco infrastructure still has a diversity problem – following years of consolidation, the sector is over-reliant on a few key players that have exercised control over the market. The sector, considered critical infrastructure, is being monopolized by the established duopoly.

Being overly reliant on just one or two suppliers with proprietary solutions is a risk to any business should there be any interruption to supply – as demonstrated by 'Rip and Replace' in the USA. A lack of suppliers also drives a lack of competition, which slows progress, fails to drive down prices and does little to spur innovation. One of the best ways to secure systems is to ensure diversity within the supply chain by adopting open, interoperable networks.

3 Five Principles of Security for an Open Future

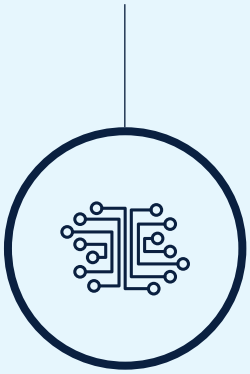
The above factors demonstrate the significant security benefits of cloud-based systems when it comes to contained software environments and the protecting against the triple threat to hardware infrastructure. However, the telco industry must not be complacent when enacting its much-needed open systems transition.

That is why Mavenir has established Five Principles of Security for an Open Future. These principles are designed to guide both Mavenir and our peers as we look to deliver a resilient future for telecoms that thrives on diversity and competition.

The Five Principles for an Open Future:

PRINCIPLE 0NE

Open architecture must drive supply diversification, preventing over-reliance on any single supplier

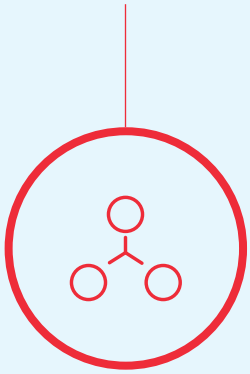


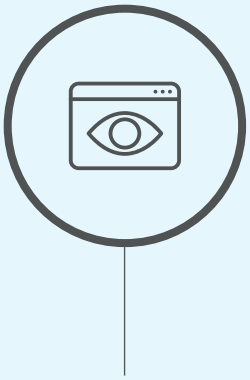
PRINCIPLE TWO

No piece of equipment or software should result in vendor lock in

PRINCIPLE THREE

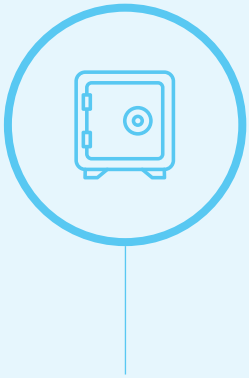
No equipment or software should compromise entire units if it needs to be replaced or upgraded





PRINCIPLE FOUR

Open, interoperable systems should provide full visibility and control of their network's endto-end security



PRINCIPLE FIVE

Open, interoperable systems should adopt a Zero Trust approach from foundation to transmission

4 The Vast Opportunity of Openness

Realizing our web3, metaverse and smart mobility vision will depend on having a robust and flexible network that underpins it – the time to lay the foundations for our future is now. Open architectures that utilize the vast resources of cloud computing offer advantages across vectors including security, but far beyond too.

Lower costs and better energy management

Running computing functions offsite is also likely to vastly reduce energy costs. The carbon footprint of cell towers has been difficult to contain, and their power management remains a challenge for tower operators and mobile network operators.

Furthermore, the energy crisis resulting from the Russian invasion of Ukraine and its impact on prices, pooling backend functions into a server farm rather than at a cell site has profound implications for better power management – details of which will be provided in a further Mavenir report in 2023.

Reduced hardware should also reduce emissions-heavy hardware production, given that research shows access infrastructure currently accounts for 60-80 percent of network's energy consumption. The digitalization of the industry is helping to make communities not only more productive, but more sustainable and resource efficient.

Accelerated speed of deployment, agility and capacity

Expansion in a remote environment is exponentially quicker. Rather than replacing hardware, network functions can be scaled up rapidly in a remote cloud environment.

The back-end scalability of data means that more users can also be added seamlessly – contributing more than 50 percent capacity to the same piece of hardware in a short period. It's also much easier to upgrade, modify, or enhance functions in an open system.

Greater interoperability and flexibility

Not being locked into hardware means mobile network operators can take a much more dynamic approach to their operations – to scale up and down as needed, change suppliers, swap in and out different elements without needing to make wholesale changes to the system.

This differs from the hardware approach which is closed, and where fixing or changing one element requires tearing out the whole system. It's like when a lightbulb goes out in your home, you simply replace the lightbulb and move on. In a closed telco ecosystem, rather than replacing the lightbulb, you have to rip out all of the wiring in the house and replace it to make it work.

Interoperability drives competition which in turn drives growth, innovation and competition for generations. Network builders are already proving that they can work together in a vibrant and diverse telecoms ecosystem – from the multivendor DISH network in the US to the burgeoning diversification of Deutsche Telekom's network in Germany.

2. Open RAN security based on zero trust architecture

Rooted in the principle of "never trust, always verify," Zero Trust is designed to protect modern digital environments by leveraging network segmentation, preventing lateral movement, providing Layer 7 threat prevention, and simplifying granular user-access control.

A zero trust architecture (ZTA) is a cybersecurity architecture that is based on zero trust principles and designed to prevent data breaches and limit internal lateral movement. The following is the relevant text from NIST publication 800-207 - 'Zero Trust Architecture'.

A "zero trust" (ZT) approach to cybersecurity is primarily focused on data and service protection but can and should be expanded to include all enterprise assets (devices, infrastructure components, applications, virtual and cloud components) and subjects (end users, applications and other nonhuman entities that request information from resources).

In this new paradigm, an enterprise must assume no implicit trust and continually analyze and evaluate the risks to its assets and business functions and then enact protections to mitigate these risks. In zero trust, these protections usually involve minimizing access to resources (such as data and compute resources and applications/ services) to only those subjects and assets identified as needing access as well as continually authenticating and authorizing the identity and security posture of each access request.

Support of a zero-trust architecture requires each O-RAN component to comply with established functionalities and protections. The guiding principles for the O-RAN allience ongoing work, includes:

1. Support integration with an external identity, credential and access management system (ICAM) using industry standard protocols
2. Require authentication and authorization on all access
3. Support role-based access control (RBAC)
4. Implement confidentiality on connections between O-RAN and external components
5. Implement integrity checking on connections between O-RAN and external components
6. Support encryption of data at rest
7. Support replay prevention
8. Implement security log generation and collection to an external security information and event management (SIEM)

THE ANALYSIS IN THE FOLLOWING SECTIONS ASSUMES A CLOUD-NATIVE OPEN RAN NETWORK WITH NETWORK FUNCTIONS MODELED AS CONTAINERIZED MICROSERVICES.

Open RAN security is built on the following tenets:

- Secured communication between Network Functions
- Secure framework for the Radio Intelligent Controller (RIC)
- Secured platform for hosting the Network Functions

3. Secured communication between network functions

This section explores following areas that relate to providing secure communication between all Network Functions in Open RAN.

a. Secure communication on all interfaces
b. Ensuring trust based authentication of communicating endpoints
c. Trusted Certificate Authorities for Identity Provisioning

3.1 Secure Communication on all interfaces

O-RAN Alliance specifies an open and secure architecture that includes secure interfaces between all its components.

Communications exchanged on these interfaces are cryptographically protected for encryption, integrity protection and replay protection.
It should be noted that several O-RAN Alliance specifications are still on-going and accordingly security work is happening in parallel. For protection of the CUS-Plane messages on Open Fronthaul LLS interface, O- RAN Alliance is currently in the process of determining all the threats and vulnerabilities, and their impact on the CUS-Plane.

3.2 Establishing trust based on mutual authentication

Mutual authentication is used for authenticating two entities with each other and setting up a secure encrypted connection between them. Mutual authentication prevents introduction of rogue NFs or xAPPs in the network.

Operator X.509 certificates are used for mutual authentication while establishing secure connections using IPsec and TLS protocols.

All network elements in an Open RAN, i.e. O-CU-CP, O-CU-UP O-DU and O-RU, support X.509 certificate- based authentication and related features such as auto-enrollment and auto-re-enrollment with an operator Certificate Authority (CA) server using a protocol such Enrollment over Secure Transport (EST) or 3GPP- specified CMPv2.

The xAPPs in the Near-RT RIC are securely on-boarded like any other microservice and the O-RAN Alliance is expected to use CA signed X.509 certificates to authenticate before communicating over the E2 interface.

Step 1-2
When the O-RU powers on, the O-CU-CP, O-CU-UP and O-DU instances that are allocated to serve that O-RU are instantiated by the orchestrator, if not already instantiated.

Step 3
An O-CU-CP, O-CU-UP and O-DU performs EST or a CMPv2-based certificate enrollment procedure in compliance with 3GPP with the CA server to obtain an operator certificate. The operator certificate is used for subsequent authentication when establishing an IPSec or a TLS connection.

Step 4
Necessary OAM actions are performed on the O-CU, if any, including changing of default passwords.

Step 5-9
Steps 5-9 are executed as part of the O-RU power-on sequence. Key security related steps are explained below:

- The O-RU obtains its IP address, the EMS or OSS address from a DHCP server using one of the DHCP options specified in O-RAN M-Plane specification section 3.1.1 and 3.1.4.
- The O-RU performs certificate enrollment procedure with the CA server to obtain an operator.
- The O-RU shall notify the EMS or OSS with a NETCONF call home. O-RU's operator certificate is used to authenticate with the EMS. OSS / EMS shall configure the O-RU with the secondary NETCONF controller's address (i.e. the address of the O-DU).
- The O-RU shall notify the O-DU with a NETCONF call home to securely obtain O-RU's configuration. O-RU's operator certificate is used to authenticate with the O-DU.

3.3 Trusted Certificate Authorities

It is recommended that the certificate authorities (CA) should be audited under the AICPA/CICA WebTrust Program for Certification Authorities.

This promotes confidence and trust in the CA servers used in Open RAN for authenticating network elements.

4. Secure framework for RIC

4.1 Security aspects of near-real-time radio intelligent controller (Near-RT RIC)

The Near-RT RIC is an SDN component that contains 3rd party extensible microservices (called xApps) that perform selected radio resource management (RRM) services for the NFs that were traditionally managed inside the gNB.

The Near-RT RIC interfaces with the O-CU-CP, O-CU-UP and the O-DU via the O-RAN standardized open E2 interface. The Near-RT RIC also interfaces with the Non-RT RIC and the service management and orchestration framework via the A1 and O1 interfaces.

4.1.1 Secure Interface between Near-RT RIC and the O-CU-CP/O-CU-UP/O-DU

Interface security is explained in § 4.2

4.1.2 Conflict resolution and xApp authentication

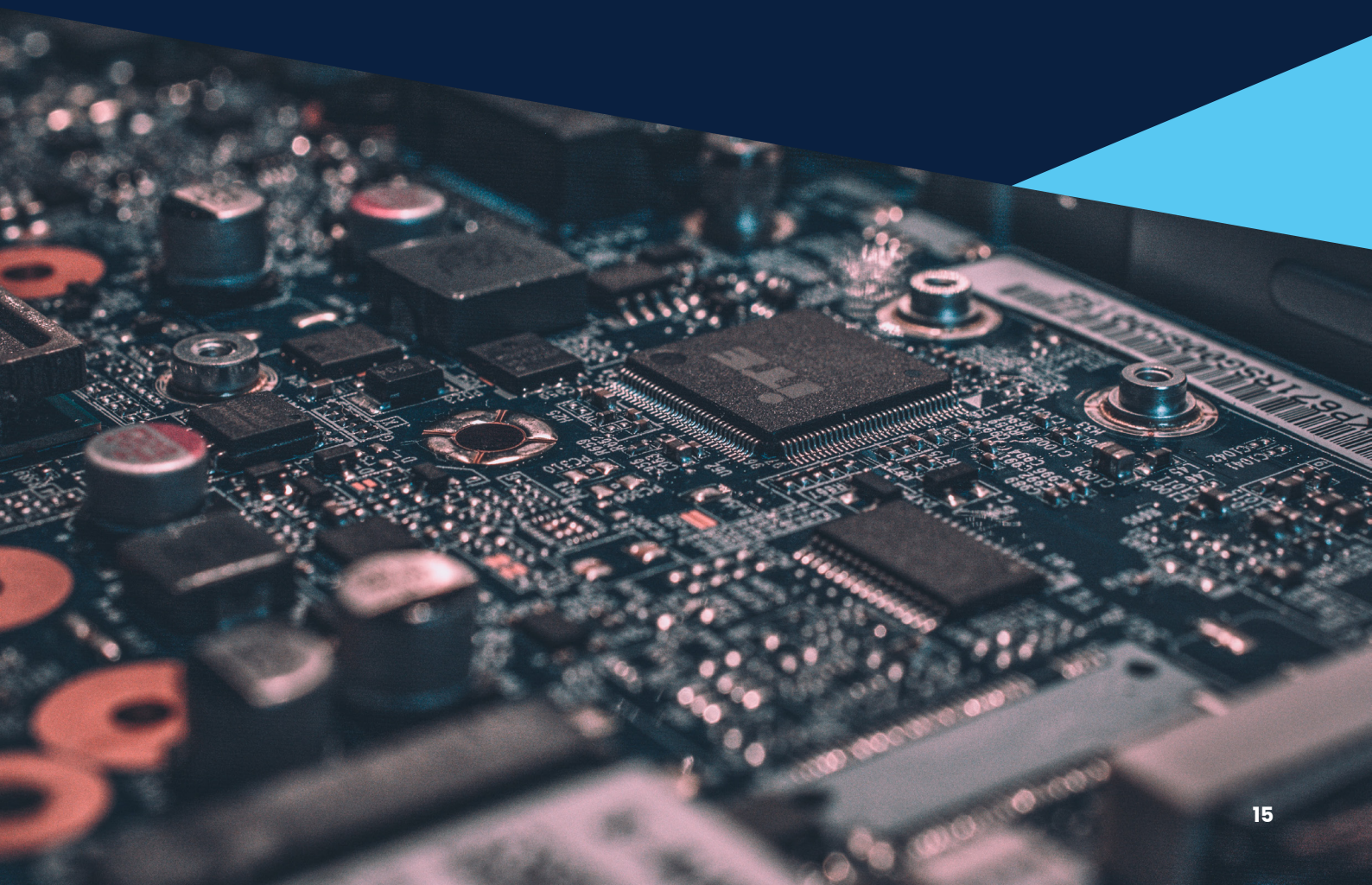
The conflict resolution among the xApps is not necessarily a security issue but can lead to vulnerabilities if not handled properly.

The key security aspects of the Near-RT RIC include:

- Secure E2 Interface between the Near-RT RIC and the O-CU-CP / O-CU-UP / O-DU
- Conflict resolution and xApp authentication
- User identification inside the Near-RT RIC
- Authorization function for xApps and A1 interface
- Secure A1 interface, O1 interface

While the xApps in the Near-RT RIC initiate the RIC subscription procedure with the E2 nodes, the subscription manager in the Near-RT RIC platform, enforces the subscription policies and keeps track of the subscriptions initiated by the xApps and the RAN functions, and event triggers associated with those subscriptions. The subscription manager can resolve signaling conflicts among the xApps by one or more of the following means:

- The subscription manager will not allow more than one xApp to subscribe to the same NF based on the same event trigger.
- If more than one xApp subscribes to the same NF and gets the same indication messages from the E2 node, then the subscription manager can allow them to simultaneously control the NF of the E2 node, as long as they do not optimize the same or closely inter-dependent parameters pertaining to the NF.
- If more than one xApp subscribes to the same NF and gets the same indication messages from the E2 node and if they optimize closely inter-dependent parameters, then the subscription manager can allow them to simultaneously control and optimize those parameters by using locks and backoff timers to retain mutual exclusivity.



4.1.3 User identification inside the Near-RT RIC

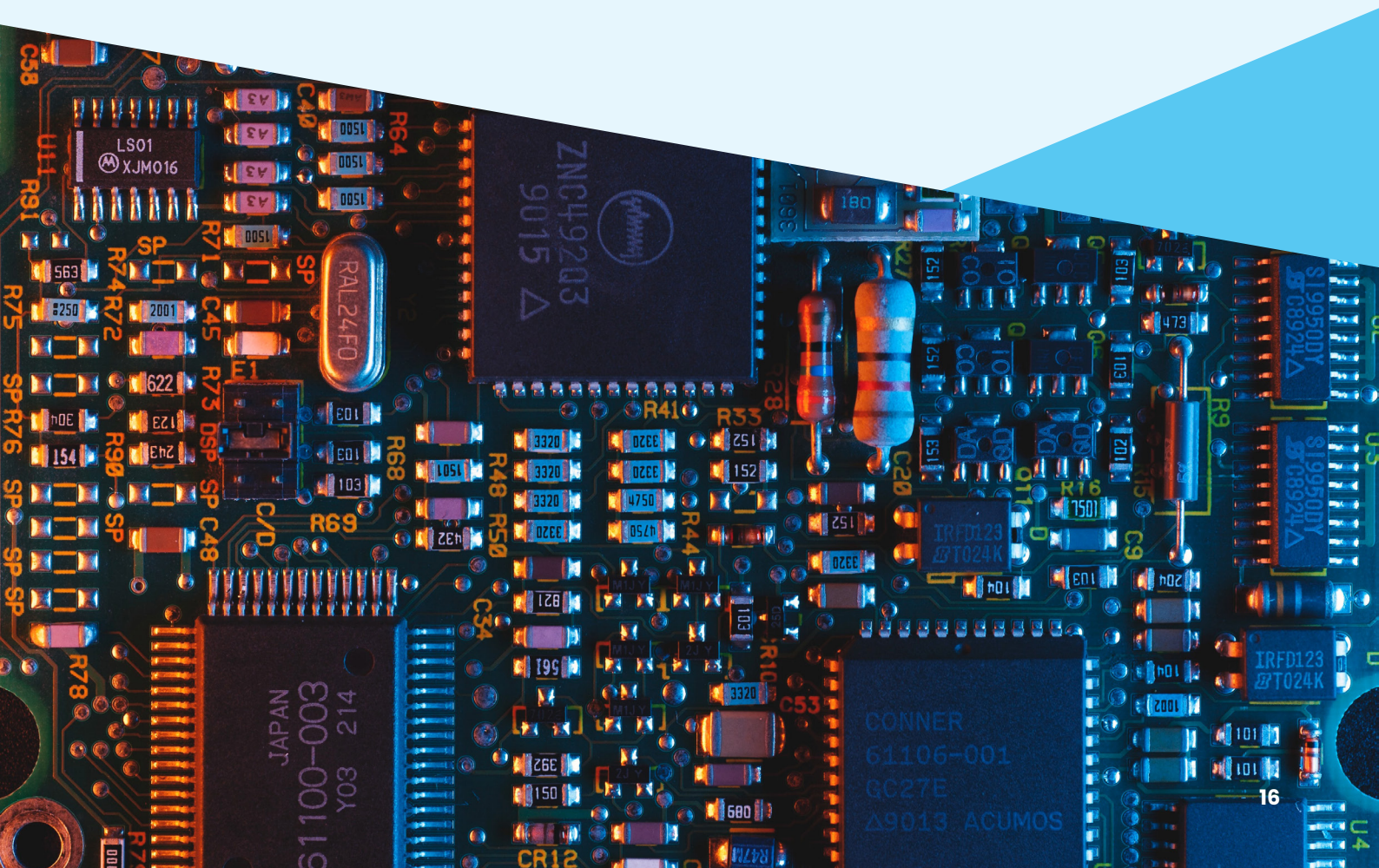
Maintaining privacy of the users is of utmost importance inside the RIC. ORAN WG3 is working on the UE identification inside the Near-RT RIC that can be addressed by a combination of 3GPP-defined Trace ID, 3GPP-defined RAN UE ID, temporary RAN network interface-specific UE IDs, and by correlating these IEs with one another.

Typically, it is ideal for the Near-RT RIC to maintain persistence of UE identification for near-RT granularities, ranging from 10 ms to 1 s. The xApps are not exposed to UE permanent ID. Invalidation of the temporary IDs in the RIC when they are released in RAN nodes will be handled via normal E2 communication. In neither case is this a UE privacy issue or a DoS attack threat.

4.2 Security aspects of Non-Real-Time Radio Intelligent Controller (Non-RT RIC)

The Non-RT RIC is a component in an O-RAN system for non-real-time control of the RAN through declarative policies and objective intents.

1. The Non-RT RIC is deployed in a service management and orchestration framework (SMO) and provides declarative policy guidance for cell-level optimization by providing the optimal configuration values for cell parameters over the O1 interface.
2. The Non-RT RIC also sends declarative policies for UE-level optimization to the Near-RT RIC via the A1 interface.
3. The Near-RT RIC then translates the recommended declarative policy from the Non-RT RIC over A1 interface into per-UE control and imperative policy over the E2 interface.
4. The Non-RT RIC develops ML/AI-driven models for policy guidance and non-RT optimization as rApp microservices. These rApps interface with the xApps over the A1 interface to optimize a set of procedures and functions in the underlying RAN.



The key security aspects of the Non-RT RIC are the following:

- Secure interface between Non-RT RIC and the O-CU-CP / O-CU-UP / O-DU
- Conflict resolution between the Non-RT RIC and the O-CU-CP / O-CU-UP / O-DU
- Authorization framework for A1 interface and rApps
- Secure A1 interface

4.2.1 Secure Interface between Non-RT RIC and the O-CU-CP / O-CU-UP / O-DU

Interface security is explained in § 4.2

4.2.2 Conflict resolution between the Non-RT RIC and the O-CU-CP / O-CU-UP / O-DU

Usually, a conflict in RRM arises when the RAN uses policies and objective intents different from the Non-RT RIC to manage the underlying RAN nodes such as the O-CU. This may be the source of rApps causing signaling conflicts with the functioning of the underlying RAN nodes. However, using the RIC subscription policies, mutual exclusivity can be enforced causing the subscribed procedures from the RAN to be managed by the Near-RT RIC, without causing signaling conflicts.

5. Secure platform for network elements

O-RAN Alliance RAN architecture is built on a fully cloud-native architecture – the same cloud architecture that is the bedrock of today's internet and public cloud. The cloud-native network functions in the O-RAN network viz. O-CU-CP, O-CU-UP, O-DU, Near-RT RIC and Non-RT RIC, are hosted on a cloud-native platform, very similar to the cloud-native platform used in the cloud computing industry. The O-RU is a PNF and thus hosted on a non-virtualized platform.

IN THE FOLLOWING SECTIONS WE TAKE A HOLISTIC LOOK AT SECURITY ASPECTS OF THESE PLATFORMS.

5.1 Secure platform for cloud-native network functions

The O-RAN architecture uses a cloud-native platform to host O-CU-CP, O-CU-UP, O-DU, Near-RT RIC and Non-RIT RIC network functions. A typical cloud-native platform typically, consists of three distinct layers:

1. Container-based application software
2. Cloud-native software stack comprising an immutable OS, Kubernetes and Container runtime
3. Cloud-native hardware infrastructure

THE FOLLOWING SECTIONS LOOK AT SECURITY FEATURES OF EACH OF THE THREE LAYERS THAT MAKE UP A CLOUD-NATIVE PLATFORM

5.1.1 Security of a container-based application software

A workload is an application or a service deployed on the cloud. Containers offer a packaging infrastructure in which applications and dependent libraries are abstracted from the environment in which they actually run.

Containers are generally perceived to offer less security than virtual machines. But it's worth noting that containers have been in use in the IT industry to build applications such as for banking which are no less critical than telecom applications in terms of security requirements, and the industry has evolved itself in automating its security and establishing best practices. The following industry standard practices are used in Open RAN to ensure security of the container-based application software:

a. Secure software development based on "secure by design" principles
b. Automating security testing based on DevSecOps
c. Vulnerability management in Open Source and 3rd party libraries

Secure software development based on "secure by design" principles

A software development life cycle (SDLC) is a framework for the process of building an application from inception to decommission.
In the past, organizations usually performed security-related activities only as part of testing—at the end of the SDLC. As a result of this late-in-the-game technique, they wouldn't find bugs, flaws, and other vulnerabilities until they were far more expensive and time-consuming to fix. Worse yet, they wouldn't find any security vulnerabilities at all.

A secure SDLC involves integrating security testing and other security-related activities into an existing development process.

Using a secure SDLC process for the workloads deployed in a O-RAN network such as xAPPs in Near-RT RIC, O-CU-CP and O-CU-UP and O-DU microservices, ensures early detection of flaws in the system, awareness of security considerations by all stakeholders involved in designing, development, testing and deployment of containers, and overall reduction of intrinsic business risks for the organization.

Automating security testing based on DevSecOps

Since the beginning of modern computing, security testing has largely been an independent activity from software development. Security focused QA professionals performed testing during the testing phase. A DevSecOps approach to the container development lifecycle ensures that security is built-in at every stage of the CI/CD pipeline.

The philosophy behind DevSecOps is to begin security testing early in the SDLC. DevSecOps integrates various security controls into the DevOps workflow such as secure coding analysis using static application security testing (SAST), automated unit, functional and integration testing. This enables developers to fix security issues in their code in near real time rather than waiting until the end of the SDLC.

O-RAN Alliance architecture software takes advantage of the advancements in 'security automation' and trend in cloud computing towards "shift left." This ensures that workloads run in the O-RAN network are validated securely (during build/deployment phase) and risk-based timely actions are taken when vulnerabilities are found before they are deployed in operator network.

Container runtime

The cloud-native software infrastructure includes a lightweight, Kubernetes-specific OCI- compliant container runtime versioned with Kubernetes such as CRI-O to reduce the risk of vulnerabilities. The cloud-native software infrastructure (container -specific OS, container runtime, disk …) must support running in FIPS mode by using FIPS 140-2 validated cryptography.

Native security with Kubernetes

Kubernetes provides several built-in security capabilities to secure the container environment including network security, resource isolation, access control, logging and auditing.

Kubernetes built-in controls that help in tightening security include:

a. Role based access control (RBAC) Use of RBAC in the cluster provides a framework for implementing the principle of least privilege for humans and applications accessing the Kubernetes API.

b. Configure the security context for pods to limit their capabilities.

c. Pod security policy sets defaults for how workloads are allowed to run in the cluster. These controls can eliminate entire classes of attacks that depend on privileged access.

d. Use Kubernetes network policies to control traffic between pods and clusters.

e. Kubernetes' network policies allow control of network access into and out of the containerized applications. In addition to this feature, software-based firewalls may be deployed to control container to container communication within or across different clusters.

f. Use namespaces to isolate sensitive workloads and create security boundaries – separating workloads into namespaces can help contain attacks and limit the impact of mistakes or destructive actions by authorized users.

g. Assess the container privileges – Adhering to the principle of least privilege and provide the minimum privileges and capabilities that would allow the container to perform its intended function.

h. Use mutual Transport Layer Security (TLS) for all inter cluster and intra cluster communications.

i. Capability to encrypt the etcd datastore to protect infrastructure and application secrets or to support integration with external vaults.

Leveraging Kubernetes operators for security

Kubernetes operators are software extensions to Kubernetes that make use of custom resources to manage services and their components in an automated way. These operators can be leveraged by the cloud-native software platform for specific security purposes:

- Hardware management operators to restrict the need for applications of elevated privileges
- Compliance operators to continuously monitor the compliance of the cluster
- File integrity monitoring operators to detect any attacks impacting the platform integrity
- Platform management operators to fight configuration drift and enforce a secure configuration by eliminating human errors
- Audit and log operators to manage the audit configuration and the log forwarding to a SIEM

A cloud-native-based O-RAN network can leverage native security controls in container runtime and container orchestration platforms such as Kubernetes, to provide defense in depth security for the containerized workload that they host.

Secure configuration of the cloud infrastructure based on industry benchmarks.

The cloud infrastructure is configured based on industry best practices such as CIS benchmarks for operating system, Docker and Kubernetes, and Network Equipment Security Assurance Scheme (NESAS) jointly defined by 3GPP and GSMA provides a consistent framework and common external audit program for multiple vendors and operators.

This ensures that appropriate security controls are put-in-place in the platform, thus reducing its attack surface.

Some of the common security controls include disabling unused ports and unused service, principle of least privileges (PLoP) for workloads, protecting data in storage, user access control using RBAC, etc.

All virtualized platforms in an O-RAN network are hardened as per 3GPP's security assurance specifications and other well-known industry benchmarks such as those from CIS. This ensures that security controls are implemented at every layer of the platform thus reducing the platform's attack surface.

Detecting and remediating configuration errors with cloud security posture management

Misconfiguration is the #1 cause of cloud-based data breaches. A mechanism is needed to make sure the configuration of the deployed cloud resources is correct and secure on day one, and that they stay that way on day two and beyond. This is referred to as cloud security posture management (CSPM).

The cloud industry has used CSPM security tools to continuously monitor cloud environments for detection of cloud misconfiguration vulnerabilities that can lead to compliance violations and data breaches.

With the adoption of a cloud-native architecture in O-RAN based networks, an operator now has the means to deploy advanced CSPM tools to guard against natural "drift" of on network configuration and reduce the potential for attacks.

Commercial cloud-native hybrid platform

Standardizing on a commercial cloud-native hybrid platform enables the operator with the following security benefits:

- A Kubernetes-certified platform with the flexibility to run securely on-prem or in a virtual private cloud, supporting O-RAN topology variations from the SMO, RICs, CUs, and DUs with zero-touch provisioning,
- Extended software lifecycle with dynamic updates that address new CVEs and optimizations over time into disconnected environments,
- Support for multi-tenancy so that multi-vendor software can be securely hosted in the same cluster,
- Support for infrastructure compliance scanning (OpenSCAP) and remediation,
- A container registry with vulnerability scanning to eliminate vulnerabilities on O-RAN platforms (e.g Near Real-Time RIC) and associated xApps and rApps

5.1.2 Security considerations with a cloud-native hardware infrastructure

O-RAN enables decoupling of hardware and software, allowing for a platform to be built from different vendors.

5.1.2.1 Secure storage of credentials and data at rest

It is recommended that O-RAN hardware comes with a hardware-based security module like TPM to manage, generate, and securely store cryptographic keys. Hardware-based security modules are also meant to provide a hardware root of trust to enable secure computing by providing a secure key storage enclave with minimal cryptographic functions primarily in the signing and signature verification space.

The data at rest must be encrypted using keys generated from hardware-based security modules.

5.1.2.2 Establishing software chain of trust

Zero-trust cannot be achieved without the full participation of all the elements in the trust chain for a network.

Trusted hardware

The hardware is built with a tamper resistant "hardware root of trust" device that provides a secure environment for storing cryptographic keys and for attestation of certificates and all the software running on that hardware. The device will expose a simple user interface for the application to use when it needs to use the device for storing keys, retrieving certificates etc.

Trusted software

Software signing is enforced at all software layers including the firmware, cloud-native software stack and container workloads at time of deployment, as well as authenticated version upgrades to make it more difficult to introduce malicious software into operator- controlled elements.

Establishing end-to-end chain of trust with secure boot

Secure boot requires that every boot up is starting from a piece of software that cannot be updated in the field. This piece of software is referred to as Core Root of Trust for Measurement (CRTM).

Thereafter, during the boot process every software program in the platform will be integrity verified before its execution by the software at the lower layer. This establishes an end-to-end software chain of trust. The trust anchor for the software integrity verification is software signing certificate. In the O-RAN network, it is recommended to use secure boot based on hardware root of trust and software signing to establish an end-to-end chain of trust.

5.2 Secure platform for O-RU

An attacker with unauthorized access to the management interface of an unprotected O-RU could allow an attacker to steal unprotected private keys, certificates, hash values and/or inject malwares and/or manipulate existing O-RU software. An attacker could further launch denial- of-service, intrusion, and replay attacks on other network elements including an O-DU.

Therefore, hardening of the O-RU platform will ensure enough equipment security to substantially reduce the attack surface that would otherwise exist in an unprotected O-RU. Security precautions on the O-RU can be divided into three aspects.

1. Supply chain security
2. Physical security
3. Network security

Supply chain security ensures that throughout the supply chain process of manufacturing, from O-RU to its final installation site and commissioning, a controlled secure chain of custody process is followed. This ensures that the O-RU is properly tracked and tagged.

Physical security ensures that the physical O-RU is sealed with non-tamper-able screws that cannot be easily broken or opened and in the event of tampering or forced opening, all O-RU functionality will be disabled so that the O-RU becomes inoperable.

This is in addition to all the physical and logical ports being secured and isolated, so that they cannot be used as a vulnerability entrance into the extended RAN network.

From a network security point of view, O-RU ensures that all authentication and communication security protocols are correctly performed and followed. To ensure reliable and secure software upgrades, the TPM procedures are implemented so that rogue software downloads are prevented.

Finally, hardening features, such as disabling unnecessary software components and interfaces when not in use, running software at the correct privilege-level, scrambling/encryption of data in storage, and secure boot and hardware-based security module, are part of the comprehensive security processes on the typical O-RU to ward off as well as prevent unauthorized access to the O-RU.

6. Conclusion

At the heart of Open RAN is the use of cloud-native architecture, the same architecture that is the bedrock of today's internet and public cloud. Security practices in virtualized deployments are mature and used across the cloud computing industry. Virtualized deployment in telecom networks is not new. Operators already have virtualized infrastructure in their data centers and many have deployed virtual workloads for other components in the network including: packet core, IMS, and other applications such as CDN. With a disaggregated architecture, operators will now additionally benefit from security expertise and experience of today's large cloud infrastructure suppliers in managing the security of large IT cloud environments.

Operator regains control as the operator now understands what is required to build and maintain a secure infrastructure. Open RAN is built on a cloud-native platform with clear responsibilities and accountability established between hardware/ infrastructure suppliers, a hybrid-cloud platform supplier, and RAN software suppliers. It enables network operators to select suppliers that meet all the required industry security standards and certifications.

Open RAN leverages several security industry best practices used in the cloud computing industry. A "shift- left" strategy in the software development process integrates security controls and practices into every phase of the software development. With DevSecOps integrated into the CI/CD pipeline, this also brings automation into secure code reviews and security testing. Use of automated tools for detection, remediation of vulnerabilities in open-source software and detection, and management of secure posture provides an operator with quick detection and resolution of anomalies in the network.

O-RAN Alliance's architecture for RAN is built on the secure foundation of zero trust where network elements mutually authenticate with each other in order to communicate. All communication between them is transported over a secure interface per industry best practices specified by O-RAN Alliance's security specifications. While standards are still evolving, the Open RAN pioneers and ecosystem vendors like Altiostar, Mavenir, Fujitsu and Red Hat, as well as early adopters like Rakuten, Vodafone, Telefonica, NTT Docomo and DISH have ensured that all the interfaces are secured using certificate based security.

Every network element in the Open RAN network undergoes platform hardening as per 3GPP's security assurance specifications and other well-known cloud computing industry benchmarks such as CIS. This protects the network from an attacker gaining unauthorized access and subjecting the network to Denial-Of- Service (DOS) attacks or gaining illegal access.

In summary, open, standardized interfaces remove vulnerabilities or risk that comes with proprietary and potentially untrusted implementation and provides an operator full visibility and control over the cloud environment and network in general.

7. Appendix

Acronyms

3GPP	3rd Generation Partnership Project
5G	5th Generation
CA	Certification Authority
CI/CD	Continuous Integration/Continuous Delivery
CIS	Center for Internet Security
CMP	Certificate Management Protocol
CNF	Cloud-native Network Function
CP	Control Plane
CPRI	Common Public Radio Interface
CRI-O	Container Runtime Interface for OCI compatible runtimes
CRMT	Core Root of Trust Measurement
CSP	Cloud Service Provider
CU	Central Unit
CUS	Control, User & Synchronization
DOS	Denial of Service
DDOS	Distributed Denial of Service
DTLS	Datagram Transport Layer Security
DU	Distributed Unit
EST	Enrollment over Secure Transport
FIPS	Federal Information Processing Standards
GSMA	Global System for Mobile Communications Association
HSM	Hardware Security Module
ICAM	Identity, Credential and Access Management
LLS	Lower Layer Split
LUKS	Linux Unified Key Setup
MAC	Mandatory Access Control
MEC	Multi-access Edge Computing
MITM	Man-in-the-Middle
NDS	Network Domain Security
NESAS	Network Equipment Security Assurance Scheme
NF	Network Function
NIST	National Institute of Standards and Technology
NR	New Radio
NR-RIC	Near Real Time RIC
OCI	Open Container Initiative
O-CU	O-RAN Central Unit
O-DU	O-RAN Distributed Unit
O-RAN	Open Radio Access Network
O-RU	O-RAN Radio Unit
PDCP	Packet Data Convergence Protocol
PNF	Physical Network Function
RAN	Radio Access Network
RBAC	Role Based Access Control
RIC	Radio Intelligent Controller
RLC	Radio Link Control
RT-RIC	Real-Time Radio Intelligent Controller
RRM	Radio Resource Management
RRU	Remote Radio Unit
SAST	Static Application Security Testing
SCRM	Supply Chain Risk Management
SDAP	Service Data Adaptation Protocol
SDLC	Software Development Life Cycle
SIEM	Security Information and Event Management
SLA	Service Level Agreement
SMO	Service Management and Orchestration
SSH	Secure Shell
STG	Security Task Group
SUCI	Subscription Concealed Identifier
TCO	Total Cost of Ownership
TLS	Transport Layer Security
TPM	Trusted Platform Module
UE	User Equipment
UP	User Plane
VNF	Virtualized Network Function
ZTA	Zero Trust Architecture

MAVENIR

For further information or enquiries, please contact

PR@mavenir.com