Towards 5G Embedded Trust: Integrating Attestation Extensions in Vertical Industries

Thanassis Giannetsos, Dimitris Papamartzivanos, Sofia Anna Menesidou, Sophia Karagiorgou
Ubitech Ltd., Digital Security & Trusted Computing Group, Greece
Email: {agiannetsos, dpapamartz, smenesidou, skaragiorgou}@ubitech.eu

I. Edge Trust Assurance Services for Cyber security Awareness in 5G-enabled Ecosystems

Recent efforts have made substantial progress towards realizing next-generation smart-connectivity "*Systems-of-Systems*" (SoS). These systems have evolved from local, standalone systems into safe and secure solutions distributed over the continuum from cyber-physical end devices, to edge servers and cloud facilities. The core pillar in such ecosystems is the establishment of a 5G infrastructure capable of managing service graph chains with embedded trust [1] comprising both resource-constrained devices, running at the edge, but also microservice technologies (e.g., Docker, LXC) [2].

Under the perspective of cloud application providers and developers, there is an increased interest in emerging mixed-criticality use cases that are apparent in a number of key sectors, from telecommunications to energy, from transport to healthcare and from robotics to military (as stated in the 5G empowering vertical industries report provided by the 5G-PPP association [3]). Such services are characterized by strict performance requirements, fast service deployment times (including also secure remote asset management), scalability and flexibility in the composition of the service graph chains as well as operational assurance but exhibit different levels of security, privacy, and trust requirements and priorities. This generates a clear trend towards decentralized architectures and business models implemented through the Mobile Edge Computing (MEC) concept (Figure 1): The available (trusted) computing resources are positioned at close proximity to the edge devices focusing on protecting the security and integrity of the generated data. Edge and fog computing nodes, mini-data centers (DCs) coexist in a 5G-enabled environment supporting the deployment of mixed-crticality services [4] positioned to execute either closer to the edge or the back-end cloud infrastructure, depending on the underlying connectivity requirements and available resources. The goal is to **enable high scalability by decomposing a mixed-criticality application into a set of "cloud-native" and "edge-running" microservices**, with different trust considerations, and managing **secure accelerated offloading capabilities** for distributing the resource intensive processes to the backend, thus, limiting the workload that needs to be managed at the edge. This will allow the overall system to reach its full potential, in a secure and trusted manner, without impeding safety.

This work was supported by the European Commission, under the ASTRID, ASSURED and PUZZLE Horizon 2020 project with Grant Agreement No. 786922, 952697 and 883540, respectively.

As a consequence, we must understand such mixed-criticality services inherently and increasingly as federated safety-critical systems operated by multiple stakeholders with different security goals. A good example can be considered in the emerging field of Intelligent Transportation Systems (ITS) and Connected Cars, where comprised Electronic Control Units (ECUs) are produced by suppliers, integrated into cars by the OEMs, car owners bring and integrated third-party devices like smartphones and everything is connected to automotive backend systems and roadside-infrastructure units via cellular communication and/or the 5G network medium. In this context, there is a high volume of generated data that must be efficiently managed and processed, by the deployed microservices, both for safety-decisions (i.e., collision avoidance) or other less safety sensitive applications such as infotainment [5]. This, however, sets the challenge ahead: *does the data stays on-board of the edge device (e.g., vehicle) or is it shared with other neighboring systems or the backend infrastructure for more efficient processing capable to achieve the current requirements of security and safety convergence?*

To this end, the 5G community is already considering technologies such as Network Functions Virtualization (NFV) & Mobile Edge Computing [6] intelligent orchestration. These enabling pillars are based on advanced virtualization capabilities and have the potential to transform existing cloud-based inftastructures into distributed datacenters. They can allow for the customization of current service graph chains to the needs of mixed-criticality applications and expose them as "network slices" [7] and, through MEC, such applications can achieve full network-awareness and zero-perceived latency. For instance, in the ITS scenario, a (trusted) "application orchestrator" will be responsible for managing the lifecycle of all deployed microservices: Essentially it will manage network acceleration mechanisms for offloading the resource intensive non-safety-critical operations from the edge to the backend infrastructure, thus, supporting the execution of a new breed of real-time, "edge-running" and latency free safety-critical services that need to operate in strict security boundaries.

In such a setting, following the decomposition of mixed-criticality applications and services into chainable components and the composition of service graphs that are placed over the virtualised programmable infrastructure, special focus has to be given into a set of threats and vulnerabilities at the

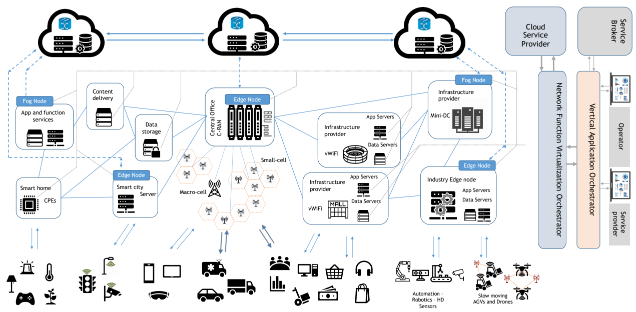

Fig. 1. 5G Enabler for Next-Generation "Systems-of-Systems" with intelligent network control and microservice deployment

component and service graph level. Assuring the security and integrity of the developed services and chainable components has to be realised during design time, deployment time, and run-time as well as a posteriori based on cybersecurity data analysis. In order to assure that the deployed mixed-criticality microservices are meeting their strict security, trust and operational assurance requirements (without impeding their safety), components must be enabled to make statements about their state (with verifiable evidence) so that other systems can align their decision towards establishing communities of trust. This goes substantially beyond current strategies where network security management services are considered in an isolated manner relying on traditional perimeter security and forensics in a "catch-and-patch" approach without dwelling on the assurance and safety of the overall network as a whole.

However, a key challenge in view of such high uncertainty processes is to establish and manage trust between microservices starting from bi-lateral interactions, between two single components, and continuing as such systems get connected to ever larger entities, thus, achieving the vision of (trust aware) service graph chains. An open question is: *How can we make sound statements on the trustworthiness of single systems and extend these trust assertion to establish dynamic chains and communities of trust across "Systems-of-Systems"?*

Compounding this issue requires strong attestation and verification services, as a means of assurance and trusted interoperability between deployed VNFs, covering all phases of a service graph execution; from the trusted boot and integrity measurement, enabling the generation of static, load-time evidence of the comprised containers correct configuration (Configuration Integrity Verification (CIV) [8]), to the run-time behavioral verification of those safety-critical components of a service graph providing strong guarantees on the execution correctness [9], thus, enhancing the scalability when composing secure service graphs from potentially insecure microservices and containers.

More specifically, we argue that what is needed is the development of virtualization-aware attestation anchors for the secure configuration, deployment, operation, orchestration and verifiable computing of edge processes, safety-critical programmable cloud native components. This needs to include novel OSS (NFVs, NFVI and VIM) trust extensions, leveraging root-of-trust capabilities (e.g., TEE-enabled middleware) that guarantees and simplifies the trust relationships between all layers in the NFV runtime stack. The aim of this process is to enable the support of extended trust aware service graph chains, for highly virtualized and software environments, with verifiable evidence on their correctness and functional safety, from their trusted launch and configuration to the runtime attestation of both behavioural and low-level concrete execution properties about an NFV's integrity.

The reason behind employing attestation mechanisms as a means of operational assurance is multifold. One of the main challenges in managing network security in today's heterogeneous and scalable infrastructures is the lack of adequate containment and sufficient trust when it comes to the behaviour of a remote system that generates and process mission critical and/or sensitive data. This high level of trustworthiness which will not only include integrity of system hardware and software but also the correctness and integrity of the generated data flows will, in turn, reduce the overall attack vector and allow for the more effective operation of the deployed microservices.

Overall, by supporting such multiple security and trust-oriented deployments, the vision of safety-critical SoS solutions can be achieved through the provision of: (i) personalized cybersecurity functions in multi-tenant environments, (ii) enhanced operational assurance and attestation capabilities by defining trust zones (that can be dynamically updated) comprised of distinct pockets of infrastructure where resources operate at the same trust level and similar safety-critical functionality, thus, minimizing the number of allowed pathways and limiting the potential for malicious threats to affect safety-critical applications, and (iii) automated orchestration of advanced security deployments in an efficient way.

References

[1] "Ericsson Mobility: 5G uptake even faster than expected." [Online]. Available: https://www.ericsson.com/en/press-releases/2019/6/ericsson-mobility-report-5g-uptake-even-faster-than-expected

[2] M. De Benedictis and A. Lioy, "Integrity verification of docker containers for a lightweight cloud environment," *Future Generation Computer Systems*, vol. 97, pp. 236–246, 2019.

[3] G. P. Architecture Working Group, "5G empowering vertical industries." [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2016/02/BROCHURE_5PPP_BAT2_PL.pdf

[4] P. Jamshidi, C. Pahl, N. C. Mendonça, J. Lewis, and S. Tilkov, "Microservices: The journey so far and challenges ahead," *IEEE Software*, vol. 35, no. 3, pp. 24–35, 2018.

[5] I. Krontiris, T. Giannetsos, P. Schoo, and F. Kargl, "Buckle-up: autonomous vehicles could face privacy bumps in the road ahead," in *18^{th} escar Europe : The World's Leading Automotive Cyber Security Conference (Konferenzveröffentlichung)*, 2020.

[6] M. . ETSI Group Specification (GS) MEC 003, version 1.1.1, "Mobile edge computing (MEC); framework and reference architecture."

[7] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: enabling enterprises' own software-defined cellular networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 146–153, 2016.

[8] B. Larsen, H. B. Debes, and T. Giannetsos, "Cloudvaults: Integrating trust extensions into system integrity verification for cloud-based environments," in *European Symposium on Research in Computer Security*. Springer, 2020, pp. 197–220.

[9] N. Koutroumpouchos, C. Ntantogian, S. Menesidou, K. Liang, P. Gouvas, C. Xenakis, and T. Giannetsos, "Secure edge computing with lightweight control-flow property-based attestation," in *2019 IEEE Conference on Network Softwarization (NetSoft)*, 2019, pp. 84–92.