FUNCTIONAL SAFETY AND SECURITY

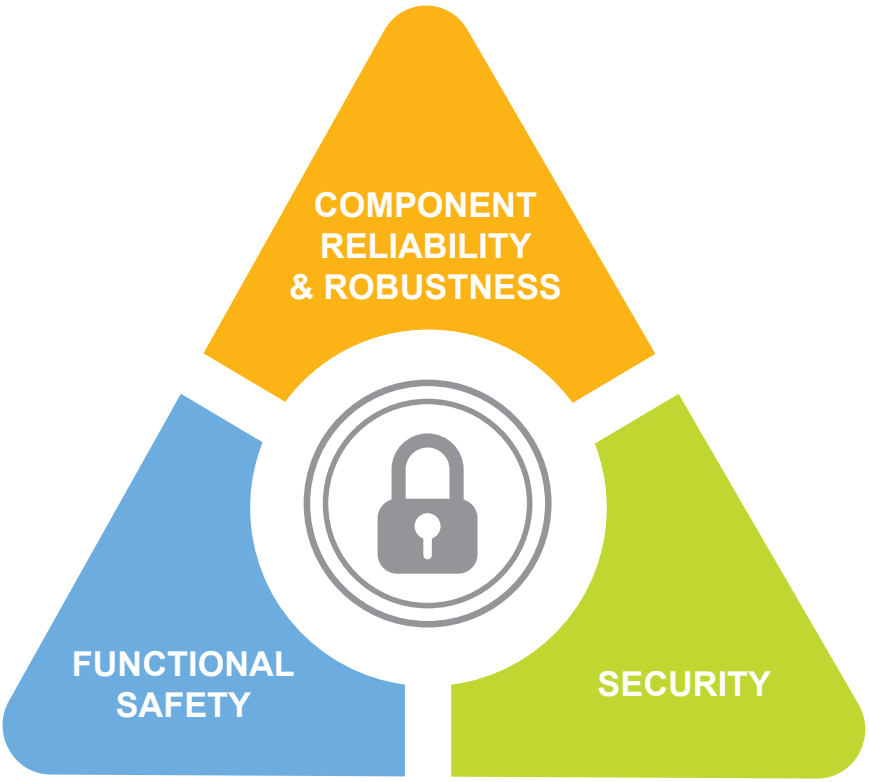
ESSENTIAL AND COMPLEMENTARY DISCIPLINES FOR MODERN SYSTEMS

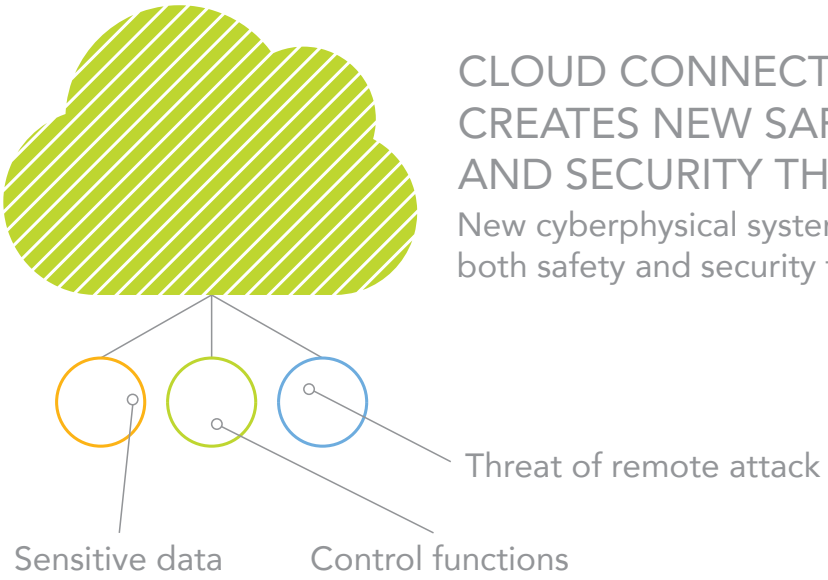
Functional safety and security are well-known and important design principles in many industries. This whitepaper provides a global introduction to and comparison of the functional safety and security domains and their individual aspects. It is intended to help executives, practitioners and experts familiar with one of the two domains to broaden their knowledge of the other and to identify synergies and interdependencies. Additional whitepapers focusing on select topics in more detail will follow this publication.

INTRODUCTION

Functional safety is a very well-established design principle in many industries. Its primary focus is to reduce injury or health risks to humans to an acceptable level when using technical equipment, but it also includes reducing the risk of damage to property or the environment. It addresses human errors, tool errors, typical technical failures, and often results in system monitors and redundant implementations that are robust in multiple ways. Obviously, this requires an end-to-end view in the assessment of these risks and for finding the appropriate solutions. Functional safety reviews always involve verification and an assessment or even a certification of the effectiveness of the measures implemented on the product and process sides.

Security is also required in many industries to protect any kind of user asset from being accessed unlawfully, one way or another. These assets are manifold and include monetary assets, intellectual property rights, or any kind of private data. Protection of integrity, authenticity and availability are other objectives in security, and here we see a strong overlap with functional safety.





CLOUD CONNECTIVITY CREATES NEW SAFETY AND SECURITY THREATS

New cyberphysical systems need both safety and security features

THE RELATIONSHIP BETWEEN FUNCTIONAL SAFETY, SECURITY AND SYSTEM SAFETY

As outlined above, functional safety aims to reduce the risk of unintended hazards, caused by malfunction of one or more system components, whereas security aims to reduce the risk of intentional threats, caused by humans. So, why have safety and security traditionally been separate disciplines? Is functional safety the same as safety?

Let's start with the first question. Security traditionally applies to systems that contain or process secrets or other sensitive data, such as personal devices and ICT systems. Functional safety traditionally applies to embedded systems that have a dedicated control function. These embedded systems are housed within larger electro-mechanical systems and they interact with the physical world around us. Due to this interaction, these embedded systems typically have strong safety requirements. A trend we see today is that these "physical" systems are increasingly connected. This connectivity enriches these systems and adds situational awareness through information exchange with nearby devices and in-the-field introduction of new features and bug fixes through over-the-air (OTA) updates. This connectivity also exposes these systems to new cyberthreats, i.e., humans may try to attack these systems remotely. There is a new, emerging category of devices with safety and security aspects. These devices are sometimes called "cyber-physical systems."

Now to the second question. It is intuitive to understand that operational safety is a must-have for physical electro-mechanical systems; the risk of harm to persons should be minimized during operation of the system. So how do functional safety and security relate to operational safety? Both help to achieve operational safety. Without them, unintentional hazards or intentional threats could negatively affect the functions implemented in the electronics of the system that control the operation of the system. Operational safety itself is only one aspect of safety at system level. Other aspects include electrical, mechanical and chemical safety. Hence, system safety typically depends on operational safety, which itself depends on functional safety if electronics and software are involved. It also often depends on security, especially if the systems are connected and more easily subject to attacks by humans. The Jeep hack that occurred a few years ago is a perfect example. The hackers managed to break into the radio. Once inside the vehicle network, they could control various safety functions, including the brakes, and thus negatively affect operational and system vehicle safety.

A much more technical and detailed comparison of functional safety and security can be found in IEC TR 63069, "Industrial-Process Measurement, Control and Automation – Framework for Functional Safety and Security." The basis for functional safety is set by IEC EN 61508, "Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems." Other standards, such as ISO 26262, "Road vehicles – Functional Safety," have built upon this original work to provide similar frameworks tailored to specific markets and applications. Similar standards are also being developed for security, including ISA/IEC 62443, "Industrial Communication Networks – Network and System Security," and ISO/SAE 21434, "Road Vehicles – Cybersecurity Engineering." Furthermore, new standards such as ISO/PAS 21448, "Safety Of The Intended Functionality (SOTIF)," are being developed to address safety at a system level.

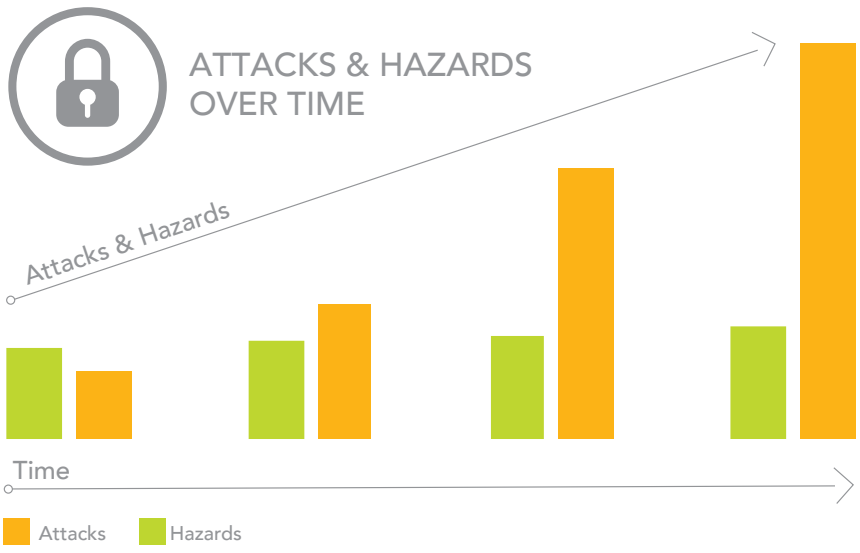
COMPARISON

In essence, in order to achieve functional safety, it is important to have a holistic end-to-end view of the system in question to understand the potential risks. Indeed, any analysis of an element or a subsystem malfunction must look at it as a standalone malfunction and as the propagation of this malfunction into the system. Ultimately, the analysis must evaluate the criticality on the overall system. Likewise, it is paramount in security to make an end-to-end assessment of all relevant ways to breach the security of the system. In a typical "hack" of a system the hacker requires a number of stepping-stones before reaching the target. The aforementioned Jeep hack is a famous example in which the car radio was the initial entry point for the attack. In the end, because of all subsystems being connected to each other, the entire system was compromised, including the car's brakes. As end-to-end assessments tend to be very complex, we will still want to partition the system as much as possible to keep the scope required for the analysis as small as possible.

In functional safety as well as in security, the analysis starts with a definition of the assets that need protection. This is a straightforward act in functional safety; it is generally the health of the people exposed to the system, possibly extended to refer to an entire environment. In security, assets may differ from case to case. In the case of an electronic wallet, it is the money contained in that wallet. For a video game manufacturer, it may be their intellectual property contained in the software. In fact, the firmware of devices often contains intellectual property that needs protecting. Hence, there is an overlap in attributes protected by the two domains. They both aim to protect integrity and availability, but security's scope is broader as it also aims at protecting confidentiality and authenticity.

The next step in security is to establish a threat model and a set of attack vectors. This is where we see substantial differences between functional safety and security. First of all, notations are somewhat different: functional safety starts with a hazard analysis and risk assessment (HARA) that is conceptually comparable to a threat model and later analyzes failure modes that are conceptually comparable to attack vectors. In the case of functional safety, there is a collection of possible user mistakes, development errors and inadvertent technical failures of the system that need to be covered. This is a more static picture and changes only slowly over time when new insights are gained, such as when new technical failure modes are discovered, new health risks are identified, or legal requirements change. The security threat model defines what type of attackers are considered. It identifies their capabilities and, in particular, how close they can get to the system, i.e., only remotely via internet, or even physically with the device in the hands of the attacker. Obviously, the assumed capabilities of the attacker have a major impact on the security solutions required. This is a major difference with functional safety for which the faults happen inadvertently and randomly during operation. In contrast, security attacks are intentional. What's potentially worse is that the attacker becomes more proficient as the attacks evolve to become stronger and cheaper over time. This is the main driving force. And it does not stop here—new attacks are found at an increasing pace all the time. The Spectre and Meltdown attacks are a famous example of when side-channel attack techniques were used creatively. These techniques were familiar to other fields, but were applied with dramatic success to modern microprocessor architectures. All this introduces an element of time and aging to any solution, and with that also the need for periodic field updates of possibly quite large parts of the software stack to maintain the level of security required. In conclusion, we can state that both domains take a risk-based approach; the risk analysis for functional safety is more quantitative in nature than the risk analysis for security which relies more on estimations.

The risk analysis for functional safety is more quantitative in nature than the risk analysis for security which relies more on estimations.



Another open point is the interaction between security and functional safety and between the respective components implementing them. Issues may be caused by unsafe or insecure component interactions, none of which may have failed and, in fact, satisfy their requirements. In particular, the reaction to a fault or incident may cause issues in other parts of the system. For example, a security subsystem that senses it may be under attack might want to reset itself and maybe even other parts of the system, including those that implement safety functions. However, it generally is not a good idea to "blindly" do so; reset may be considered a safe and secure state, but it doesn't necessarily mean that none of the safety goals are violated. Such issues could be identified using existing hazard analysis methods, provided that there are cross-checks between security and functional safety concepts. Other issues caused by unsafe or insecure interactions of components may only be found in operation as "unknown unknowns." It is for that reason that there are initiatives that investigate whether existing methodologies, possibly from outside of the two domains, could be used to identify and eliminate or mitigate such issues early in the development process. One idea is to apply system theoretic process analysis (STPA) which, unlike the traditional hazard analysis methods, can be started in early concept analysis to assist in identifying safety and security requirements and constraints. The idea here is to design safety and security into the larger system architecture and design, starting with a global analysis in the early phases, and refining it in later phases when the system design is refined and more detailed design decisions are made.

Issues may be caused by unsafe or insecure component interactions, none of which may have failed and, in fact, satisfy their requirements.

The solutions to meet these design principles are usually cast in safety/security functions that need to be implemented. The substantial overlap of security and safety functions is one of the most obvious targets for synergy between the two disciplines. An analysis of acceptable remaining risks is also needed, as there is always a residual remaining risk no matter what precautions are taken. In security, the acceptable remaining risk is often expressed as a minimum number of points every attack needs to score in a suitable attack rating metric. In functional safety residual risks are evaluated with a failure mode effect and diagnostics analysis (FMEDA), which results in a probabilistic metric of hardware random fault (PMHF), checked against a target value defined by the necessary safety integrity level (SIL).

This brings us to the next commonality: in functional safety, as well as in security, it is good practice to perform a verification of the claimed properties. This is done with a set of tests that are as thorough and complete as possible performed during the design stage, but also later during the production of the product. At advanced levels these tests will also involve formal proof of the claimed properties.

Both domains apply similar roles and process steps. For example, both domains require concepts and architectures to be defined by experienced architects to prevent common mistakes, reviews to be performed by independent reviewers and assessments by security and safety assessors to confirm that functional safety and security goals have been achieved. Both domains require expert know-how and experience with a typically steep and long learning curve; however, the expertise and experience is different for the two disciplines. Still, someone with good knowledge and expertise in one discipline will be able to get started more quickly in the other discipline, though they still will not immediately be an expert. Therefore, training is essential for both domains. Typically, a combination of classroom and on-the-job training is offered to grow talent. Furthermore, we see a growing consensus that a single team or expert cannot efficiently handle both domains in parallel. The trend is therefore to establish close links between both domains and expert teams, and to have regular interaction and alignment between them throughout both processes. Safety goals may be inputs to security process, and security countermeasures may need to be validated by functional safety experts in close and iterative loops. Lastly, organizations need the right mindset to continuously improve awareness and to effectively address both domains. This mindset should stem from the organization leaders to help build and nurture the safety and security culture.

Lastly, organizations need the right mindset to continuously improve awareness and to effectively address both domains. This mindset should stem from the organization leaders to help build and nurture the safety and security culture.

When we look at regulation, then the picture is quite different between (functional) safety and security. Safety regulation is in place for many markets – including aviation, railways, road transportation and medical applications – and compliance is typically enforced through type approval. Currently, legal requirements are often poorly developed for security, although there are initiatives for select markets. One such example is that the World Forum for Harmonization of Vehicle Regulations (UNECE WP. 29) is developing a new International Whole Vehicle Type Approval (IWVTA) scheme that includes security regulation. Under this scheme, which Europe plans to adopt and enforce in 2022/2024, a certified cybersecurity management system (CSMS) becomes a prerequisite for vehicle manufacturers (OEMs) to achieve type approval for new vehicles. Furthermore, suppliers will be required to provide evidence of compliance to the OEM.

Compliance is a good and sometimes necessary step, but may not be sufficient. In both disciplines, it is easy to make claims, but it's less easy to validate those claims. It is therefore common practice for both design principles to give customers more independent assurances of the quality and effectiveness of all the implemented measures. Firstly, an organization can be certified by an accredited third-party for compliance with engineering standards. This process certification is the major trend in the automotive industry and the intent of the ISO 26262 and ISO/SAE 21434 standards. We expect the same trend in the industrial market (IACS) with IEC 61508 and ISA/IEC 62443, and possibly in other markets, too. Secondly, product certification can be applied to provide assurance that the products fulfills its security and/or safety objectives. This can be applied for functional safety. However, self-assessment is also allowed by the standards, provided that the assessor within the own organization has sufficient levels of independence and expertise. In larger companies, which usually have centralized functional safety teams with sufficient expertise and experience, the practice is to move away from product certification in favor of performing confirmation measures in-house.

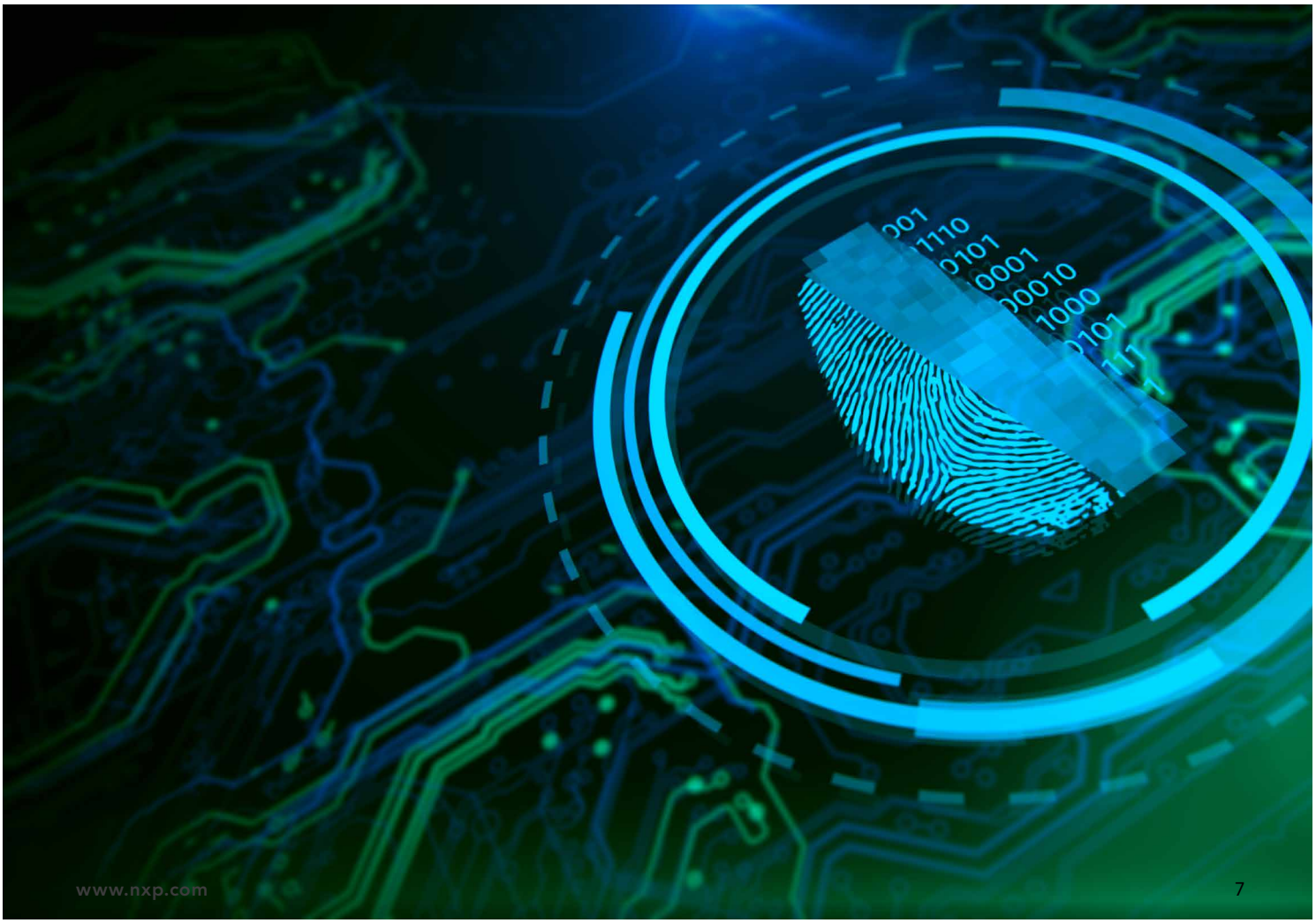
There are currently a few industries that require product certification for security. The most notable examples are the banking industry with EMVCo standards and countries with the Common Criteria standard (e.g., France). Also, some areas of governmental interest, such as travel documents, passports, ID cards, etc., often require a Common Criteria certification with assurance levels defined in the EAL metrics. Yet, for the much larger field of IoT, certification of security claims made for products are still in their infancy. It could, however, be expected that this will change over time as security becomes more and more critical for virtually every infrastructure. The privacy laws now enforced in Europe are a first example of this trend. Arm® has recently launched a proprietary PSA certification scheme for subsystems used in IoT products. NXP founded an industry consortium that created the SESIP evaluation methodology/certification scheme which addresses an even larger and more comprehensive part of the system stack in IoT products than PSA does.

Despite concept, design, implementation and confirmation reviews by independent experts throughout the development phase, possibly extended by process and product certification, incidents with products in the field can never be fully excluded. This is the case for both disciplines, and in both cases, there is usually pressure to quickly identify the root cause and define and implement mitigation strategies and solutions to address these, as the real-life impact may be significant. Furthermore, it is useful to monitor devices in the field for anomalies. It is for that reason that a security operations center (SOC), a concept stemming from the information technology (IT) world, is now also being considered and, in some cases, already implemented for operational technologies (OT) such as connected vehicles and industrial systems.

Monitoring is also being considered for safety, especially in the context of SOTIF to detect for unknown unknowns. However, there are also differences; security has to deal with an evolving landscape, in which new threats and attacks occur regularly and where existing ones become easier and cheaper to mount. It is for that reason that organizations should also implement threat monitoring to stay ahead of the curve.

CONCLUSION

There is a clear need for companies to master both disciplines and drive alignment between them. The similarities between functional safety and security are surprisingly strong. This should help anybody with good knowledge and expertise in one area to get a head start in the other field. There are also many potential opportunities for synergies due to these similarities. However, there are also significant differences that can be spotted once taking a closer look at the various aspects of the two domains. The main difference lies in the fact that security is a more dynamic field, with new attacks being invented all the time, and at an increasing pace. Also, the malicious intent of an attacker and their dedication to find weaknesses and exploit them has a huge impact on the nature of the security solutions that are required. By the same token, this also means that security is a rather fast-moving target. Any proposed solution needs to be revisited and possibly updated and verified on a regular basis. Also, because of the ever-increasing complexity and sheer number of connected products within a single ecosystem, the scope of such an analysis keeps changing and widening. To this end, appropriate mechanisms need to be in place, quite possibly involving more than one company along the value chain of a product. However, the autonomous driving challenge will lead to a more dynamic environment



from a safety perspective as well. The drastic increase of artificial intelligence usage in the vehicle and the potential for collecting a huge amount of information through the cloud will generate a need for regular safety critical application software updates. This involves a move from a hardware-based safety solution that is quite "static" to a "hardware as a platform" safety approach allowing a flexible but still safe software safety concept. As such, the safety integrity will have to be ensured all along the supply chain: development, verification, update, deployment, etc. Following this autonomous path also mandates an evolution from fail-safe to fail-operational systems. This leads to an even bigger challenge for functional safety and security: how to keep the system running after a fault or an attack?

Although security certification is still in its infancy for many emerging markets such as IoT, it is strongly believed that a meaningful and independent security certification will be needed for those markets to flourish sustainably. This concept has been repeatedly proven to be successful in many industries already, not just for security but also for (functional) safety. On the organizational process level, this requires engineering standards tailored to specific verticals, such as ISA/IEC 62443 for industrial and ISO/SAE 21434 for automotive. This will require new comprehensive evaluation methodologies/certification schemes like SESIP that allow for a more flexible and modular approach to product (re)certifications.

Both functional safety and security are already very complex disciplines in their own right. As industries increasingly require both, it is therefore paramount to make these two disciplines compatible and remove complexities and contradictions. This will require aligning methodologies, metrics, processes, concepts, architectures, countermeasures, and, last but not least, close alignment and interaction between disciplines and teams. Given all the commonalities outlined above, this should be an ambitious but achievable target; it's also a huge opportunity.

Companies that are strong in both fields are in an ideal pole position to be leading players in the connected, technically complex world of tomorrow, playing to their strengths and exploiting the synergies of these two design principles to the fullest, making sure they complement each other in an integrated fashion.

www.nxp.com/safeassure

Document Number: SAFETYSECWPA4 REV 0