Sanctus: An Architecture for Trusted Products

Malcolm Shore
Canterbury University

Sherali Zeadally
University of Kentucky

Andy Clark
Royal Holloway College

Abstract: The last two decades have seen a fundamental shift in the manufacturing, sourcing and operation of technology, which has raised concerns in state security agencies about the cybersecurity risk to government and critical infrastructure. Sophisticated cyber attacks continue to be launched by state actors worldwide, while the engineering practices in common use have failed to deliver a commensurate improvement in technology cyber security. Cyber attacks continue to be successful against commercial networks, leading the US Government to encourage government agencies to look towards models such as zero-trust networking and tailored trustworthy spaces. There has been progress in product engineering, with formal methodologies such as Correctness by Construction (CbyC) successfully producing commercial products with increased trustworthiness. However, the adoption of these techniques has been limited, and governments are now increasingly resorting to an approach of technology Balkanization, where import and use of products and components may be restricted based on their country of origin. Even in the early stages of this strategy, the effect upon the economy is significantly adverse. We propose an alternative to technology Balkanization by combining trustworthy engineering approaches with the use of a national security component we call a sanctum which together can deliver sovereign trust.

Keywords: Cybersecurity, Balkanization, trust

Introduction

The evolution of technology over the last two decades has been rapid, particularly in the telecommunications field. Simple internet services became the World Wide Web and have evolved into sophisticated cloud technologies; the local network, which allowed terminal access to connected servers, has evolved into the internet of everything (Chandhok, 2014). The bricks and mortar of society are continuing to give way through digital transformation into the smart cities and businesses of the future (Matt, Benlian & Hess, 2015; Chanias &

Hess, 2016). In pace with these changes, nation states and criminals have found ways to subvert technology for their own benefit at substantial cost to the rest of the world (Anderson *et al.*, 2013). Out of simple viruses and worms designed to be mere annoyances have emerged vast botnets of compromised zombie computers capable of launching devastating attacks on unsuspecting targets anywhere in the world. Simple malware attachments have evolved into sophisticated techniques, such as those used by the Platinum Group to attack computers even when they are powered off (Mimoso, 2017). One consequence of this evolution of malicious attacks is that trust in technology has plummeted. Concerns over the vulnerability of internet-connected systems continue to form the mainstream driver for cybersecurity, with increasingly sophisticated attack technologies being used by criminals and nation states (Alcaraz & Zeadally, 2015).

The threat to ICT systems has not gone unnoticed, and information security practices and standards have evolved over the years. The original United Kingdom's Department of Industry Code of Practice PD0003 was adopted as the British Standards Institute's BS7799 and, subsequently, by the International Standards Organization into what is now known as ISO27000: Code of Practice for Information Security Management System (ISO/IEC, n.d.). The US National Institute of Standards and Technology issued Special Publication 800-53: Security and Privacy Controls for Federal Systems and Organizations and more recently published the NIST Cybersecurity Framework, which adopts both ISO27000 and SP800-53 controls into a single framework for cybersecurity. Other recommendations such as the UK Cyber Essentials have been proposed, but did not achieve global recognition.

Recent attacks on the Ukrainian power grid, attributed by the Ukrainian authorities to Russia, have highlighted the continuing risks to critical infrastructure (Park, Summers & Walstrom, 2017). While much of the risk can be attributed to configuration weaknesses, technology flaws and inadequate operational management, nation-state subversion of the supply chain has been highlighted by, and become something of a hobby horse for, security agencies in certain countries. As national critical infrastructures become dependent upon more advanced computing and telecommunications technologies, some of which is sourced from countries considered to be potential adversaries, so distrust of technology becomes an increasingly significant issue and governments' responses become a major risk to the global economy (National Journal, 2018).

Various governments have attempted to develop a robust approach to technology trust. In the 1960s, the US Department of Defense introduced a set of trusted systems criteria in what was known as the Orange Book (US DoD, 1983), offering trust at levels from C2 through to A1. (C2 is the accepted entry level of assurance; A1 is the highest level applied to classified systems). The UK Government introduced an alternative scheme called the IT Security

Evaluation Criteria (ITSec), which decoupled security functionality from its level of assurance. While ITSec was a significant step forward in systems assurance, improvements were needed (Gehrke, Pfitzmann & Rannenberg, 1992). Eventually, in the late 1990s, the Orange Book and ITSec approaches merged into a single set of criteria recognized by the US, UK, Canada, Australia and New Zealand (a grouping referred to as the Five Eyes). This scheme, known as the Common Criteria (n.d.), is now recognized by 28 countries as the means of approving equipment for use by governments and national infrastructure.

The US has for many years controlled the export of advanced technologies (Clark, 2015), both military and those which may be sold for peaceful purposes but have the ability for dual use in military systems. These controls have had mixed success. Many countries in the 1990s agreed to limit the spread of a key technology – namely, encryption. Encryption systems were included as a category of strategic arms, with export controls being applied to the more powerful cryptographic products. These controls proved to be counter-productive, encouraging many countries to develop their own products in competition with products from, and outside the control of, the US. Furthermore, with the posting and exchange of high-grade cryptographic techniques and tools on the internet, any control over cryptography is now ineffective.

The US and other countries have increasingly focused on technology support for the fighter, and the US, in particular, has evolved its military strategy based on having information dominance (Miller, 2019). This requires that the US has the most advanced technologies and information-enabling and disabling systems in the world and is able to access more sophisticated microelectronics than its adversaries (Chappell, 2017). As China becomes increasingly capable in developing advanced technologies, and comes to the threshold of potential technology dominance, it becomes more difficult for the US to maintain information superiority. This is particularly serious for the United States as it increasingly sources military technology components offshore (NDIA, 2017). While the United States has traditionally had significant control over technology used globally, its dependence upon foreign components for military products puts the United States at increasing risk of technology blockade or subversion should military action involve countries of supply.

The Race to Balkanize

The need for a global framework for technology trust was addressed in the late 1990s with the establishment of the Common Criteria scheme and is based on a framework of increasingly trustworthy levels of technology evaluation. A similar approach was more recently introduced in the Cybersecurity Law issued in China in 2017 (Ning & Wu, 2017). However, this approach now appears to be inadequate for the US and Australia, both of which have suggested they cannot achieve sufficient assurance to enable Chinese vendors to participate in their 5G networks.

Telecommunications has become the first major sector of national infrastructure in which serious attention has been given to supply chain security risks. An example of this is the Australian implementation of telecommunications sector security reforms, in which carriers are required to notify the Government of any substantive network changes and use of certain vendors' technologies may be limited or banned. The US has also taken similar measures with respect to sourcing from China and Russia, such as banning products from being used by US Government agencies (Volz, 2017). Surprisingly, this new development has not so far extended to the use of non-Chinese technology manufactured in China that suffers from exactly the same risk.

The US has also expanded the scope of its technology export control justified through the sanctions process. In 2017, the US banned the export of technology components to ZTE, leaving them unable to continue operations. Only a late reversal of this ban after payment of fines enabled ZTE to survive.

Taken together, the emerging approach of banning use of certain technologies and blocking exports is potentially the start of a slide towards what can be characterized as "strategic technology Balkanization", in which the technology used in a country will be limited to that manufactured within its geopolitical bloc.

There is a downside to any country completely or partially blocking certain advanced technologies:

- Balkanizing technology in a global manufacturing environment means repatriating much of the nation's offshore technology manufacturing. The Organization for Economic Co-operation and Development has warned nations against the associated strategy of localization, noting that this would jeopardize the benefits individual users and businesses enjoy from integrating global communications and the digital economy (OECD, 2016). More concerning, as explained by Apple to the US Government, the US has neither the facilities nor the indigenous skills to do the manufacturing (Worstall, 2013).
- In the event a class of technology is blocked, there may need to be some alternative source. This is the case in the US for semiconductors that was highlighted in the National Defense Industry Association reports (NDIA, 2017). Semiconductor foundries have a limited life, after which new foundries have to be built to support the more advanced chipsets. The cost to the US Government to ensure its semiconductor industry remains active is substantial, with an advanced foundry costing around $10B-$15B.

- Not using the most advanced technology in its infrastructure may result in a nation being restricted to what over time will become a "second-world" legacy infrastructure. This is particularly concerning for those countries with strategies based on digital transformation for which legacy technology cannot deliver the required products and services. Without enhancing that technology, these countries will be unable to compete globally, resulting directly and indirectly in an adverse economic impact (Qiang, Rossotto & Kimura, 2009). Mühleisen argues that restricting the use of technologies or legislating against them is not beneficial (Mühleisen, 2018). He recommends the development of smart policies that can optimize the benefits of new technology.
- In terms of a military strategy, and as was the case with encryption controls, the use of Balkanization may have entirely the opposite result to what is intended. The victims – adversaries amongst them – may choose to redouble their efforts to develop their own advanced technologies and become self-sufficient, which would merely compound the problem.

Regardless, the US and Australia have adopted a technology Balkanization strategy and have banned all 5G mobile network technology that is coming out of China. In doing so, they have accepted the cost to their economies both directly, from more expensive networks, and indirectly, with the potential for trade repercussions (Letts, 2019). This delays their digital transformation and results in it being delivered with more expensive and less advanced network technology. In Australia, the Government's decision to ban 5G from Chinese manufacturer Huawei has led to Optus delaying its 5G roll out and TPG cancelling plans to build a 5G network. Technology innovation in China will continue to accelerate, leaving Balkanized nations even further behind the rest of the world in delivering digital transformation.

China for its part is also pursuing a form of Balkanization. It is pursuing two key initiatives: Made in China 2025 and Internet Plus. Action by the US in blocking component exports has underlined the need for China to even more aggressively pursue independence in technology, and increases the awareness in other countries of the downside risk of using US technology.

Trust – an Alternative to Balkanization

Trusted computing

One alternative to technology Balkanization is to develop an approach to technology that can be trusted regardless of its source. Trusted technology concepts were introduced in the Trusted Computing Base (TCB) books published by the US DoD in the 1960s, the most popular of which was the Orange Book (US DoD, 1983), which covered trusted operating systems. However, the adoption of TCB to ensure trustworthy computing in the military fell victim to the economic imperatives of commercial-off-the-shelf (COTS) solutions and the few trusted systems that were developed have long since disappeared.

Despite the demise of TCB, the concepts of technology trust have continued to evolve. A basic foundation for technology trust is to design components to be secure, and to have a verification process to confirm their implementation is true to design. Secure design has attracted a significant amount of research, while verification has been addressed in industry with schemes such as Common Criteria evaluation, and more effectively through initiatives such as the Huawei Deep Evaluation Cell in the UK (Katwala, 2019). These approaches, however, offer only assurance at a point in time and do not address the issue of in-service trustworthiness.

Existing literature on trust

Marsh formalizes trust as a computational concept (Marsh, 1994) and notes that the formalism would not only enable network nodes to reason with and about trust, but would also provide network managers with another way to assess their networks – a remarkable insight into what is now a critical problem. Marsh defines basic, general and situational trust and includes such concepts as blind trust, optimistic trust that will never decrease, pessimistic trust that will never increase, and distrust where past actions influence current trust. He argues that the concepts of blind trust and permanent distrust should be discarded, as they do not belong in a rational decision-making system. Situational trust reflects the idea that different agents may calculate trust for the same entity differently, depending on their situation – and this may change as the situation changes. Marsh introduces the idea of *utility*, where an agent seeks to maximize the utility of a node for economic benefit, a far-sighted view of how trust needs to be balanced with economic gain. An interesting view from Marsh is that trust is not transitive, contrary to the views of later researchers such as Grandison & Sloman (2000). Marsh notes the problems that can occur in real-world trust: trust is a subjective phenomenon and humans use trust in a fashion clouded by emotions, wants, needs, and so forth; that there is a need to assess the rationality of agents making trust decisions and there is no *a priori* reason to assume agents are always rational. However, he does provide rules that a rational trusting entity, human or automated, should follow and provides a formal trust model in terms of calculating situational trust and co-operation thresholds.

In their survey of trust in Internet applications, Grandison & Sloman (2000) explore the properties of trust relationships and note that trust is never absolute but operates within limits. They cover the issue of infrastructure trust, i.e. the trust in the workstation being used, the local network and the network servers, by referring to the Orange Book. Their conclusion in surveying trust is that trust is the belief that an entity will act dependably, securely and reliably within a specified context; that trust can change over time; and that trust management enables information to be collected in order to make trust decisions. These concepts are as relevant today as when they were published in 2000.

Saadi *et al.* (2011) propose a trust meta-model to enable heterogeneous trust management systems to interoperate using mediators, allowing the development of composite trust models. Their model relates to technical aspects of stakeholder trust within different system models, but the meta-model can be widely applied to the more generic issues of trust. Their model consists of three elements: 1) trust roles, abstract representations of stakeholder behaviour; 2) trust relations between stakeholders in the model; and 3) trust assessment to compute the trustworthiness of stakeholders. The model includes direct and indirect trust relationships, with indirect trust reflecting the transitive trust concept referred to by Grandison and Sloman but also including the concept of reputation-based trust.

Networks have been a specific focus area for trust modelling, and in particular mobile ad-hoc networks. Jaydep Sen (2010) has proposed a framework for distributed trust management in mobile ad-hoc networks that approaches the problem from the perspective of key distribution and misbehaviour detection. Nodes in the network are considered to be cooperative, malicious or selfish and detection of uncooperative behaviour can be calculated using node-based reputation scores. Sen notes that external methods of preventing attack cannot be used when a node may be compromised, as it may be operating within a security envelope provided by network encryption. Instead, an internal method is proposed in which every node in the network monitors the behaviour of its neighbours and, if any abnormal action is detected, it invokes an algorithm to determine whether the suspected node is indeed malicious. The framework is designed to handle a range of behaviours such as dropping adverse feedback, selective broadcast and packet dropping, and tampering.

A particular focus for network trust has been Byzantine attacks against packet forwarding – a specific case of the more generic issue of supply-chain malware insertion. A Byzantine attack is one in which one or more nodes in a network may exhibit malicious behaviour. Zouridaki *et al.* (2007) propose a Hermes scheme and propose improvements to its robustness to Byzantine attacks (Zouridaki, Mark & Hejmo, 2007). The Hermes network scheme combines first-hand information on the behaviour of neighbour nodes, and second-hand reputational information passed from other nodes. Their improvements include a punishment policy to discourage selfish behaviour. The trust measurement assesses the number of correctly forwarded packets relative to the number of incorrectly forwarded packets and has an associated confidence factor. Hermes includes the concept of opinion, to generalize the idea of trustworthiness to non-neighbouring nodes. Han, Ravindran & Jensen (2007) propose a gossip-based mechanism for information exchange that is robust against Byzantine attacks for denying and faking messages – sometimes called a black hole attack. Their research indicates that, with relatively few rounds of gossip, the mechanism is robust in the presence of Byzantine attacks. Goyal & Sharma (2014) provided a short survey of Byzantine attacks in mesh networks, reporting the classification of attack types and identifying some key papers. Another survey was carried out in 2015 by Sindhuja, Nasrinbanu & Elavarasi (2015) addressing malicious node detection in data fusion sensor networks. Geetha & Sreenath (2016) also survey Byzantine attacks on the routing protocols used in mobile ad-hoc networks. They identify a number of Byzantine attacks including black hole, sinkhole, wormhole, gray, flood rushing, selfish and overlay network attacks. They identify a range of mitigations, including trust-based, incentive-based, cryptography-based, and analytical approaches. Eschenauer, Gligor & Baras (2002) consider Byzantine attacks in the context of mobile ad-hoc networks which may not have a fixed infrastructure available, and propose the use of swarm intelligence for trust distribution. Zhang (2017) presents the Byzantine defense problem from a contemporary cloud-based carrier network perspective, with a set of adversaries that range from the curious to state actors. He presents a Cybersecurity 3.0 model, which has to deal with intrusions within the Observe-Orient-Decide-Act (OODA) approach, and with a resilient multi-tiered control hierarchy of protection, detection, response and recovery. This approach mitigates attacks launched from the outside inwards that may subvert a node.

A significant component of technology trust is the implementation of the technology. In their Manifesto for High Integrity Software, Croxford & Chapman (2005) report that the user community and the software industry have been driven to accept that software defects are inevitable. This is less so in industries operating safety critical systems. These range from aircraft fly-by-wire control systems and railway signalling systems, to software in medical devices and traffic lights. In developing safety critical software, there are three key questions: what are the hazards presented to safety by the software; what can software engineers do to reduce hazards to an acceptable level; and how can the developed system be safety certified? There have over the years been various standards for safety critical systems, including the general standard IEC61508: Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems (2010) and the now withdrawn IEC60880: Software for Computers in the Safety Systems of Nuclear Powers Stations (1986). The CANDU Computer Systems Engineering Centre of Excellence of the Atomic Energy of Canada has published standard CE-1001-STD: Standard for Developing Safety Critical Software (CANDU, 1999).

Teikari & Nevalainen (2014) provide a good summary of safety critical software standards, and specifically review a number of IEC standards. They identify a number of deficiencies, one of which is that security is inadequately covered, and there is limited reference to coding standards. However, IEC 62645 (2014) does address the requirements for security, specifically the prevention of, detection of and reaction to malicious cyber acts that could lead to an unsafe situation. These include malicious modifications affecting system integrity, malicious interference with information, data or resources, and malicious changes to hardware, firmware or software at the programmable logic controllers.

A significant approach to delivery of contemporary trustworthy technology has been published by MITRE in their Cyber Resiliency Engineering Framework (Bodeau *et al.*, 2012). The framework provides a defence-in-depth approach to the development of architectural resilience practices to address the cyber threat.

In the area of semiconductors, the DARPA SPADE programme demonstrates the ability to disaggregate trust and enable the co-existence of multi-source commercial semiconductor capabilities (Chappell, 2017). Specifically, SPADE is designed to address the risk of malicious insertion through using secure parts to monitor commercial components packaged together into a single ASIC. Other strategies include authentication at any stage in the supply chain, reverse engineering to verify the design, and disaggregation into functional subcomponents.

More informal concepts around zero trust have been published by industry. One such example is the Palo Alto zero-trust approach to network security (Palo Alto, 2014). Zero Trust in this context is a data-centric network design that puts micro-perimeters around specific data or assets to allow more-granular rules to be enforced. Zero Trust networks solve the "flat network" problem that helps attackers move undetected inside corporate networks so they can find and exfiltrate sensitive data, and is often implemented using network segmentation.

In 2011, the United States National Cybersecurity Centre issued a strategic plan to develop a trustworthy cyberspace (NSTC, 2011). The objective of this plan was to mitigate strategic cyberspace vulnerabilities and ensure that the United States gains the most it can from the evolving use of cyberspace. The plan revolved around deep research into the root causes of the cyberspace problem, to develop scientific foundations and maximize the research impact, to induce change and to accelerate the transition of research into practice. The plan included priority areas of designed-in security, tailored trustworthy spaces, making the system a moving target for cyberattack, and cyber economic incentives. Noting that the absence of mechanisms to establish trust has made cyberspace vulnerable to illicit exploitations, the research theme for Tailored Trustworthy Spaces (TTS) aims to provide flexible, adaptive, distributed trust environments that can support functional and policy requirements arising from a wide spectrum of activities in the face of an evolving range of threats.

Beyond trust

Beyond trust, other efforts have been focusing on making networks immune to cyber attack. Cyber immunity is a relatively new research field of advanced anomaly detection, which incorporates automated response to attacks. In his research, Wlodarczak (2017) describes a cyber immune system as a detection/response/recovery technology that is inspired by the human biological immune system (and aligns well with three of the five categories of controls in the NIST Cybersecurity Framework). Traditional firewall and intrusion detection systems using signature schemes often struggle to detect zero day attacks, but a cyber immune system is designed to look for symptoms of the attack and provide a defense mechanism, which can contain and eradicate any form of attack exhibiting these symptoms. The key to cyber immunity is good detection through predefined genetic rules (innate immunity) or through learning (adaptive immunity), having a low-to-zero level of false positives (known in the biological sense as autoimmunity), and an effective response for anything that is detected. It is likely that a healthy system would contain many different cyber immunity modules, each focused on a specific class of attacks. Wlodarczak suggests that this would be achieved using artificial intelligence and neural network techniques combined with machine learning.

Tailored trustworthy spaces

The NSTC Strategic Plan (NSTC, 2011) defines a tailored trustworthy space as a technology domain that "provides flexible, adaptive, distributed trust environments that can support functional and policy requirements arising from a wide spectrum of activities in the face of an evolving range of threats". In other words, it is a cyberspace environment that provides a user with confidence in its security, using automated mechanisms to adjust the level of security based on the user's context to address an evolving range of threats. It can also be an isolated collection of devices, services, policies and data that interact securely, reliably, and with privacy. For the purposes of this paper, we take a product-centric viewpoint of tailored trustworthy spaces, such that a trustworthy space is the security domain within the product that can be trusted by users to maintain information security and operational integrity, using automated mechanisms to deliver cyber immunity services throughout the product.

The Sherwood Applied Business Security Architecture, or SABSA for short (Sherwood, Clark & Lynas, 2005), can be used to describe trustworthy technology spaces in terms of its broader concept of security domains. A SABSA domain is a complete description of a business space, in which there are people, processes and technology, all of which must be trustworthy according to the policy of the domain. SABSA describes the modelling of isolated and independent (interacting) domains and their associated inter-domain associations. This methodology allows for security attributes to be described within the domain and on the interacting links. The key elements of a security domain are that: its boundary is explicit; it has a common security policy; it has a security domain authority responsible for setting and ensuring the effectiveness of that policy; it interacts with other domains through a domain gateway; and the domain is responsible for enforcing its own security policy at the gateway. Domains can exist within a higher level domain known as the super domain, and are then subordinate domains, more commonly called subdomains. Subdomains inherit policy from their super domains, and may interpret it within the context of their own domain. The concept of SABSA domains can be applied to the enterprise as a whole, but also to describe security domains internal to an ICT system. Consequently, SABSA is a useful tool for describing tailored trustworthy spaces in the concept of products.

Early work on designing tailored trustworthy spaces that address in-service attacks was carried out for the US Department of Energy (Speicher, 2011). The focus was securing smart grid control systems with their underlying IP infrastructure using a combination of services which together form what is termed the *security fabric* of the TTS architecture. An important aspect of the security fabric framework is the use of secure silicon in addition to standard firmware-based management services, an approach now common in mobile devices with their embedded Trusted Execution Environment. Secure silicon provides for sensitive storage and processing as well as the trusted monitoring required in trustworthy spaces. The DoE design introduced the concept of a service-oriented architecture and a policy-driven managing device, which handled device communications with agents in other subordinate devices.

Main research contributions of this work

We summarise the main contributions of this work as follows:

- We introduce a trustworthy technology framework based on the Canadian CE-1001-STD standard for safety critical systems and the MITRE Cyber Resiliency Engineering Framework, which addresses the issues of national security at each stage from requirements capture to advanced operational capability.
- We introduce the concept of a component that we call a *sanctus,* a sovereign component which provides the trusted tailored technology space to use in product design.

- We demonstrate the application of our model to the design of more resilient smart grid systems.

Proposed Model of Sovereign Technology Trust

Starting with safety critical systems

Safety critical systems focus on the integrity of software and, while they do not have a focus on nation security, they do provide an insight into the engineering techniques that are required to deliver technology trust. The CANDU standard CE-1001-STD is an IEC-aligned practical example of the application of safety critical concepts and provides a model of engineering in six stages. This is shown in Table 1 with the associated tasks and outputs for each stage.

Table 1. CE-1001-STD Safety Critical Software Engineering

Stage	Activity	Documentation
Concept	Business Analysis	Design Input Document
Requirements Definition	Requirements Definition	Software Requirements Specification
	Requirements Review	Requirements Review Report
Design	Design	Software Design Description
	Design Review	Design Review Report
Code Implementation	Coding	Source Code
	Code Review	Code Review Report
Testing	Unit Testing	Unit Testing Procedures Unit Test Report
	Integration Testing	Integration Test procedures Integration Test Report
	Validation Testing	Validation Test procedures Validation Test Report
Verification	Hazards Analysis	Hazards Analysis Report
	Reliability Qualification	Reliability Qualification Report

Reviews are carried out through the Requirements, Design, and Implement stages. Testing validates engineering of the design functionality. Hazards Analysis, which gets its input from the Design Input Documentation, Software Requirements Specification, Software Design Description and Source Code, is intended to identify any input conditions or subsystem failures that could lead to the software shifting into an unsafe state.

Requirements Definition includes the identification of failure modes, and establishing requirements for fault tolerance and graceful degradation. The Design requires the identification of self-checks to enhance robustness to hardware failures or other system level hazards. Coding is required to defend against detectable run time errors such as buffer overflows.

Reliability Qualification involves defining and explicitly identifying the basis for a reliability hypothesis, and then defining tests that simulate the software's usage profile in order to provide evidence that the probability of failure of the software is sufficiently small for it to meet its reliability requirements.

With this six-stage process, the software used in nuclear energy generation is assured to be safe to an acceptable level.

Trustworthy technology framework

In developing our trustworthy technology framework, we have adopted the practical standard CE-1001-STD, aligned it with the MITRE Cyber Resiliency Engineering Framework, and further developed it with new architectural, design, development, and operational concepts to enable its use to address national security trustworthiness. This revised framework is shown in Table 2.

Table 2. Enhanced Framework for Technology Trustworthiness

Stage	Activity	Output/Outcome
Concept	Business Analysis	Design Input Document
Requirements Definition	Quality Driven Requirements Capture and Analysis	Software Requirements Specification
	Requirements Review	Requirements Review Report
Design	Zero Trust Design	Software Design Description
	Design Review	Design Review Report
Development	Secure Software Engineering	Source Code
	Code Review	Code Review Report
Testing	Unit Testing	Unit Testing Procedures (including security) Unit Test Report
	Integration Testing	Integration Test procedures (including security) Integration Test Report
	Validation Testing	Validation Test procedures (including security) Validation Test Report
Verification	Hazards Analysis	Hazards Analysis Report
	Reliability Qualification	Reliability Qualification Report
	Common Criteria Evaluation	CC Certificate
	Deep Security Evaluation	National Security Endorsement
Operate	Known attack detection	Real time blocking/alerting
	Resiliency	Real time response to ensure continuous operation
	Anomalous attack detection using national security algorithms	System heartbeat monitoring Real time blocking/alerting
Resilience	Byzantine attack detection	Isolation of malicious components
	Defence-in-Depth	Ensure no single control point of failure
	Cyber immunity	Auto response to, and recovery from, attack
	Survivability	Shutdown of non-essential functions

The second stage of Requirements Definition becomes Quality Driven Requirements Capture and Analysis to ensure that the specification of the technology requirements is correct and provides a solid foundation for delivering a product fit-for-purpose. This is a critical step to minimize the amount of rework required for the product to achieve full user acceptance. The MITRE Cyber Resiliency Engineering Framework provides a set of practices, which can be represented as SABSA attributes, to ensure that we can deliver security requirements down through design and implementation. These are:

- Adaptive response, taking actions to respond to an attack based on its characteristics;
- Analytic monitoring, gathering and analyzing data continuously;
- Co-ordinated Defence, managing multiple distinct mechanisms to respond to attack;
- Deception, actions to confuse or misdirect an attacker;
- Diversity, using different technologies to limit the spread of an attack;
- Dynamic positioning, to dynamically relocate elements of the system;
- Dynamic representation, to support situational awareness;
- Non-Persistence, to defeat known-location attacks;
- Privilege-Restriction, to make it difficult for an attacker to gain escalated privileges;
- Realignment, to reduce the attack surface;
- Redundancy, to avoid single points of failure;
- Segmentation, to control access to sensitive resources;
- Substantiated integrity, to ensure that critical elements of the system have not been corrupted; and
- Unpredictability, to make gaining a foothold difficult for an attacker.

We use the SABSA security domain concepts to achieve design of trustworthy spaces. The principles of Quality Function Deployment (QFD) can be applied within SABSA, starting with requirements capture and continuing through the framework, to create products and solutions that faithfully deliver the full spectrum of stakeholder requirements. QFD is one of the recognized foundations of design for trustworthy software (Jayaswal & Patton, 2006). Applying the SABSA practices is the first step in preparing to deliver trustworthy software by design.

Designing against a concept of zero trust enables security to be maintained even when the product is deployed into an environment in which there is no trust. Zero trust design requires that data is independently protected, and micro-segmentation is used to protect sensitive subsystems. Going further, segmentation should ensure that any sensitive data storage and processing is carried out in a trustworthy technology space. This then minimizes the scope of proof or evaluation for security assurance.

An important enhancement to CE-1001-STD is to adopt a practice of Secure Software Engineering rather than a more generic approach to coding to minimize the opportunity for flaws to be introduced during the development phase. By minimizing defects during development of the software, a great deal of post-testing recoding can be avoided, increasing reliability and reducing implementation costs. Using a rigorous design and implementation approach, such as Correctness by Construction (Kourie & Watson, 2012), for critical functions provides high productivity and low defects; and the associated use of SPARK/Ada for developing the code enables formal verification of correctness. This is made practical by designing the security critical functions to be in a tailored trustworthy space, which can then be targeted for rigorous engineering.

Our enhanced framework introduces the concept of independent testing. Common Criteria Evaluation validates security claims using the global government-recognized evaluation scheme, and we suggest that this continues to be an adequate approach for many environments. Where it is not, then Deep Security Evaluation goes further by incorporating source code review checking for both implementation weaknesses and the existence of malicious code. Taken together, these test regimes contribute the independent verification necessary to validate vendor claims of trustworthy technology and to confirm product integrity.

It is not sufficient to take trust at the point of launch as assuring the product for whole of life, as products are updated and environments change over time. In particular, products operating in hostile environments, or which can be remotely reached from a hostile environment, are susceptible to in-use compromise. Our enhanced framework therefore includes operational monitoring to maintain a level of trust through the operational use of the product. This involves known attack detection, typically using some form of signature matching, and anomaly detection to detect unusual and suspicious behaviour. The more advanced anomaly detection systems are able to learn what normal network behaviour looks like in order to more effectively detect anomalies. These capabilities are available for deploying as network solutions, but the techniques can also be applied within the product internal design.

Our enhanced framework addresses the need for superior resilience through introducing the MITRE Defence in Depth model, and three advanced activities from the research domain that provide through-life capability to detect issues harmful to the system and recover from them.

- Byzantine attack monitoring will address the issue of latent and stealthy malware introduced during manufacturing or later in the supply chain. While there has been significant research in this area, it has yet to appear in commercial products. A form of Byzantine attack detection could be applied in a trustworthy space to identify anomalous behaviour outside that space.

- From the MITRE Cyber Resiliency Engineering Framework, we adopt the four top level goals of Anticipate, Withstand, Recover, and Evolve to enhance the SABSA Defence in Depth model, which is supported by controls that apply the MITRE resiliency engineering objectives of Understand, Prepare, Prevent, Continue, Constrain, Reconstitute, Transform and Re-Architect. This can be seen in MITRE's diagram of the goals (top) and objectives (bottom) as shown in Figure 1.

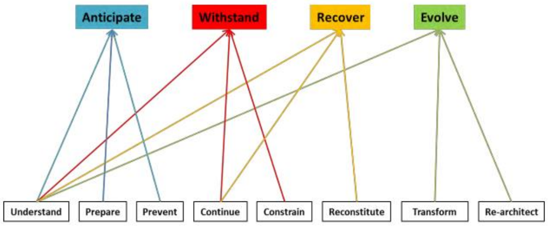

Figure 1. MITRE Cyber Resiliency Engineering Framework Goals and Objectives

- Signature-based quarantine techniques will evolve to more sophisticated advanced cyber immunity capability, which includes response mechanisms that affect not only shutdown of the attack but also "healing" of any damage done by the attack. Some progress in this area has occurred, with operating systems incorporating self-monitoring and service recovery – the first steps in the path to cyber immunity.

- Survivability, which requires the ability to fall back to a core set of critical activities in the event of overload or attack. This has not been common in industrial solutions, but can be seen in carrier mobile networks, where priority calls will be serviced even in congested networks by dropping non-priority calls. Applying the concepts of survivability to functions within a product or solution will improve the reliability of critical technology systems.

An important vector for critical infrastructure attack is the support and maintenance process, which is used to introduce changes to software in order to correct defects and provide new product features through code updates. Vendor product support is often provided remotely by a foreign national, and may require privileged access to the infrastructure. Attacks can occur by a malicious engineer uploading malware through legitimate access, or by a user-applied update being compromised. An example of the latter form of attack was the malicious code found in the CCleaner product (Collins & Hautala, 2017). The rigour applied during initial design and implementation needs to be applied for all subsequent changes to minimize the opportunity for this vector to be used, and operational monitoring will provide defence in depth.

A proposed model of sovereign technology trust

While Common Criteria evaluation was intended to deliver sufficient assurance to allow deployment of products into national infrastructures, the scheme has not been sufficient to satisfy sovereign security requirements. Despite Deep Security Evaluation facilities operating in UK, Germany and Canada, the US and Australian Governments (Varghese, 2019) continue to have insufficient trust in the world's leading technologies emerging from China.

A new model of technology trust is required to enable the deployment of globally sourced commercial products into national infrastructures. Based on our Cybersecurity Trust Framework, we propose an architectural approach to designing trustworthy technology that allows a sovereign trust module to be incorporated in a commercial product. This will then allow sovereign control of all security-related aspects of the product so that it can be deployed with assurance into the national infrastructure. We achieve this by isolating critical security information and functions to a specific domain or set of domains, which can be designed as the tailored trustworthy space within the product. The remainder of the product software can be untrusted. This is a logical extension of the design that we can see in contemporary mobile devices, where the Trusted Execution Environment allows critical security functions to be isolated in a trusted area inside the chip.

By designing the tailored trustworthy space as a discrete hardware component with secure, internationally standardized interfaces, a nation will be able to provide sovereign national security algorithms and secure key storage in a trustworthy module, which we call a *sanctus*. The sanctus could be a plug-in module of some form, either externally by the user or internally during national localisation of the product.

The sanctus as a component of the larger product would also need to be developed using our cybersecurity trust framework. It would need as a whole to be rigorously engineered in order to achieve a sovereign level of trust that can then be extended into any product in which it is used. As the secure heart of a product, it would need to have as small an attack surface as possible and ideally be formally proved.

In operational use, there would be three significant outcomes for the sanctus:

- **Secure the Information Flow**. A trustworthy product needs to be able to ensure the confidentiality and the integrity of information passing through it. The sanctus will need to have the ability to take control of interfaces so that any incoming information can be protected prior to being passed into the untrusted domain within the product. This is similar to the way in which a mobile trusted execution environment can take control of the keyboard interface for PIN entry.
- **Ensure Integrity**. Integrity is another key issue for effective security, and this means knowing exactly what software is running in the product. We can do this by demonstrating binary equivalence, meaning that the operational software matches a validated software release. By using a sanctus, securely loaded software signatures can be checked at start-up and during operation against the running code in a product. This ensures that the software has not been tampered with, and that only validated versions of software can run.
- **Monitor System Health**. For critical infrastructure, availability is often as important as confidentiality and integrity. The sanctus can be used to run real time checks of the operational state of the product and report back to a health monitoring system using a secure heartbeat mechanism. This will include the basic and advanced cybersecurity monitoring of information flows within the product and network flows touching the product and reporting any alerts via the heartbeat. These are likely to include nationally sensitive algorithms, and as such will be loaded into the sanctus rather than in the product itself. By doing this through the sanctus, an attacker cannot send spoofed reporting of health while the system is under attack or disabled.

Cybersecurity monitoring is complex. Contemporary cyber security products use one or all of signatures, learning schemes, and algorithms to deliver effective security. New cyber attacks are emerging all the time, and traditional anti-virus solutions require regular signature updates. Anomaly detection systems such as DarkTrace (2019) incorporate mechanisms to learn what normal network activity looks like, and to detect any deviations from the learned algorithms. Solutions such as Microsoft's CloudApp (2019) allows anomaly detection policies to be incorporated as algorithms. The sanctus will need to be able to support all these capabilities, as well as any advanced sovereign resilience features that have been developed.

Smart Metering Case Study

Smart grid security challenges

In this section, we apply our proposed architectural approach to the problem of securing distribution in a smart grid to demonstrate that it effectively addresses the known security challenges.

The major work on smart grid security emerged in 2010, in which smart grids are shown in a conceptual model covering generation, support systems, transmission, and distribution controlled via an operational centre, and being managed through a market approach in which service providers address the needs of customers. The business requirements for a power grid include reliability and resilience, self-healing against disruption events, and that it provides safe and efficient energy delivery (Amin, 2011). The customer impact of even a limited power loss event can be catastrophic. For example, in 2007, Mercury Energy deliberately disconnected power to a house in which a woman was on life support, resulting in her death (Henderson, 2007). There are also privacy concerns that must be addressed in smart grids – data sent regarding power usage can indicate the absence of a householder and be used to target houses for burglaries – and the data must be handled accordingly (Zeadally *et al.*, 2013).

In time of conflict or political tension, the smart grid could be an early target of critical infrastructure attack (Anderson & Fuloria, 2011). In past conflicts power and communications utilities have been targeted through air attacks or sabotage. The use of smart grids substantially reduces the cost and risk to the attacker by enabling the attack to be conducted through remote computer exploit. A significant amount of work has been carried out into identifying and mitigating smart grid threats (Otuoze, Mustafa & Larik, 2018); however, Alcaraz & Zeadally (2015) note that utilities typically have little experience of defending themselves against capable motivated cyber adversaries. The BlackEnergy cyber attack in 2015 (Lipovsky & Cherepanov, 2016) provided ample demonstration of the ability of threat actors to execute a denial of service and achieve social disruption.

Skopik *et al.* (2012) provide an insight into specific smart grid threats and vulnerabilities. They report that a smart grid system involves three tiers: the uplink from the smart meter; the backhaul to the application; and the smart grid application itself. Tier 1 attacks include local hardware and firmware manipulation and exploitation, potentially remotely, of design and implementation. Tier 2 attacks include network sniffing from the home or neighbourhood network, large scale meter takeover via malware spreading peer to peer, and backhaul concentrator node attacks. Tier 3 attacks are web attacks focusing on consumer and management services. The most likely attack from a national security perspective would come via a remote attack on the central operations management system, as the network should be fully protected against end-device penetration even should an attack on a home or neighbourhood network device be successful. However, meter misreporting of overloads could result in partial network shutdown. From a power operator perspective, the impact of these attacks falls into the category of power theft by manipulating recording or changing usage data, data theft, unit Denial of Service (DoS) by causing a meter to malfunction, or grid DoS by interfering with concentrators or mass meter compromise.

Alcaraz & Zeadally (2015) provide a further perspective on threats related to SCADA systems, including man-in-the-middle attacks, which can inject or modify control messages. These kinds of attacks can have a broad impact across the system, and are particularly significant for SCADA protocols, such as Modbus, that operate in plain text. The number of attacks on industrial control systems is increasing, and the key one for national security is denial of service. Their research indicates a number of technical approaches to achieving a Denial of Service, such as jamming of mobile components, flooding attacks, selective forwarding attacks, impersonation attacks, dropping and redirecting messages.

The attack vectors for each tier of the smart grid solution can be summarized as shown in Table 3, with the relevant adverse outcomes marked with an asterisk for the various types of attack.

Table 3. Network Attack Vectors

Tier	Attack	Power Theft	Data Theft	Unit DoS	Grid DoS
1. Water Meter	Hardware manipulation	*		*	
	Firmware manipulation	*	*	*	
	Impersonation/meter emulation	*			
	Exploitation of design weaknesses	*	*	*	
2. Utility	Network sniffing		*		
	Large scale meter takeover				*
	Message injection/modification	*	*		*
	Message redirect/drop	*			*
	Message Flooding				*
	Attacks on concentrator nodes		*		*
3. Web and Backend Applications	System penetration				*
	Theft of metering data	*			

To combat these threats, Otuoze, Mustafa & Larik (2018) have identified the key conceptual attributes relevant to ensure the smart grid is protected: authentic, available, reliable, confidential, integrity-assured, efficient, accessible, authenticated, robust, flexible, and resilient. To this we can add private, as identified by Zeadally *et al.* (2013), and the system itself needs to be trustworthy. It also needs to be as immune as possible to attacks, i.e. to have the capability to detect and trigger an immunity response. This provides the foundation for a quality driven design using a sanctus to enable zero trust throughout the smart grid.

Tier 1

Hardware manipulation of the meter is in general outside the scope of a sanctus and would need to be achieved using traditional tamper-proofing. However, designing the meter to incorporate a sanctus could address the remaining two methods of meter attack. Firmware manipulation could be detected using a sanctus through monitoring of the software integrity signature. Impersonation can be addressed as a Byzantine attack by using the sanctus to validate peer device signatures. Exploitation of design weakness would fall into either hardware or firmware exploitation.

Tier 2

The use of a negotiated encryption protocol using trusted encryption code in the sanctus would enable nationally-trusted encrypted communications to avoid data breach from network sniffing attacks. Large-scale meter takeover from viral malware moving laterally across the grid can be mitigated using monitoring modules in the sanctus. Message injection and dropping can be mitigated using Byzantine detection in the sanctus, and message modification through a man in the middle can be addressed by encrypting messages. Message flooding can be mitigated using an immunity mechanism in the sanctus that detects and reacts to a node flooding the grid by shutting it down. Concentrator nodes can be designed in a trustworthy manner in order to be resilient to attack.

Tier 3

Protecting the central smart-grid operations system and its back-end databases from attack is the most complex part of protecting the smart grid. The smart-grid software would be developed with security code and sensitive data in the sanctus, and using rigorous security engineering to ensure the code is trustworthy. Operation of the system would be designed to authenticate any critical command with a sanctus signature to ensure only sanctus sourced actions are taken. Operational monitoring would include the sanctus running the full range of configuration integrity checks across the server, automated account and access checks, and full anti-malware monitoring. Advanced resilience features might include a survival mode, which locks down the smart grid to a fixed safe state in the event of attack. Advanced security monitoring in the sanctus could prevent unauthorised access to the system to steal metering data. By having the security code and sensitive data protected in the sanctus, any compromise of the server would have limited impact.

Conclusion

In this paper we suggest that the move to technology Balkanization is a less than ideal solution to the problem of sovereignty in a global technology environment. We provide a survey of the literature on technology trust.

From this, we propose an alternative to technology Balkanization based on the Canadian CE-1001-STD safety-critical engineering process. This process is further developed to enable its application to be extended to address national security, using an architecture that incorporates a standardised sovereign component, called a sanctus, to ensure confidentiality, integrity and availability of sensitive data, algorithms, and processing. Using a smart grid case study, we have shown how this can address the known attacks in smart grids.

Further research is required to develop and trial a sanctus, and, in particular, to develop a robust interface standard that could be presented for international adoption.

References

Alcaraz, C., & Zeadally, S. (2015). Critical infrastructure protection: Requirements and challenges for the 21st century, *International Journal of Critical Infrastructure Protection*, *8*, 53-66, January. DOI: 10.1016/j.ijcip.2014.12.002

Amin, S. M. (2011). Smart Grid: Overview, Issues and Opportunities. Advances and Challenges in Sensing, Modeling, Simulation, Optimization and Control, *European Journal of Control*, *5*(6), 547-567. DOI: 10.3166/EJC.17.547–567.

Anderson, R., & Fuloria, S. (2011). *Smart meter security: a survey*. Available at: https://www.cl.cam.ac.uk/~rja14/Papers/JSAC-draft.pdf. Accessed 26 March 2019.

Anderson, R., Barton, C., Böhme, R., Clayton, R., van Eeten, M. J. G., Levi, M., Moore, T., & Savage, S. (2013). Measuring the Cost of Cybercrime. In: Böhme, R. (ed), *The Economics of Information Security and Privacy*, Berlin: Springer. Available at: https://www.researchgate.net/publication/263605690_Measuring_the_Cost_of_Cybercrime. Accessed 19 March 2019. DOI: 10.1007/978-3-642-39498-0_12

Bodeau, D. J., Graubart, R. D., Picciotto, J., & McQuaid, R. (2012). Cyber Resiliency Engineering Framework, *MITRE Technical Papers*. Available at: https://www.mitre.org/publications/technical-papers/cyber-resiliency-engineering-framework. Accessed 26 July 2019.

CANDU. (1999). Standard CE-1001-STD, Revision 2, Standard for Software Engineering of Safety Critical Software, CANDU Computer Systems Engineering Centre of Excellence, December. Available at: http://fm.csl.sri.com/VeriSure2015/talks/CE-1001-STD.pdf. Accessed 23 March 2019.

Chandhok, R. (2014). The Internet of Everything, *2014 IEEE Hot Chips 26 Symposium*, 1-29. DOI: 10.1109/HOTCHIPS.2014.7478826

Chanias, S., & Hess, T. (2016). Understanding Digital Transformation Strategy Formation: Insights from Europe's Automotive Industry, *20th Pacific Asia Conference on Information Systems*, Chiayi, Taiwan, June 2016.

Chappell, W. (2017). A Technology-Enabled New Trust Approach, DARPA presentation. Available at: https://www.darpa.mil/attachments/1TheCaseforSecureASICs_Slides.pdf. Accessed 26 March 2019.

Clark, D. (2015). U.S. Agencies Block Technology Exports for Supercomputer in China, *The Wall Street Journal*, 9 April. Available at: https://www.wsj.com/articles/u-s-agencies-block-technology-exports-for-supercomputer-in-china-1428561987. Accessed 23 March 2019.

Collins, K., & Hautala, L. (2017). Hackers hid malicious code in popular CCleaner software, *CNET*, 19 September. Available at: https://www.cnet.com/news/hackers-hid-malicious-code-in-popular-ccleaner-software/. Accessed 26 March 2019.

Common Criteria. (no date). Common Criteria Web Portal. Available at: https://www.commoncriteriaportal.org/. Accessed 27 August 2019.

Croxford, M., & Chapman, R. (2005). Correctness by Construction: A Manifesto for High-Integrity Software, *Crosstalk: Journal of Defense Software Engineering*, December. Available at: https://pdfs.semanticscholar.org/3516/e1cdf840fe0e697aa43dd6be9ab9de71120a.pdf. Accessed 27 August 2019.

DarkTrace. (2019). The Enterprise Immune System. Available at: https://www.darktrace.com/en/products/enterprise/. Accessed 27 August 2019

Eschenauer, L., Gligor, V., & Baras, J. (2002). On Trust Establishment in Mobile Ad-Hoc Networks, *Lecture Notes in Computer Science 2845*, September. DOI: 10.1007/978-3-540-39871-4_6

Geetha, A., & Sreenath, N. (2016). Byzantine Attacks and its Security measures in Mobile Adhoc Networks, *International Journal of Computing, Communications & Instrumentation Engineering*, *3*(1), 42-47. DOI: 10.15242/IJCCIE.AE0116013

Gehrke, M., Pfitzmann, A., & Rannenberg, K. (1992). Information Technology Security Evaluation Criteria (ITSEC) – a Contribution to Vulnerability?, *Proceedings of the 12th IFIP World Computer Congress on Education and Society* - Information Processing 92, *II*. Available at: https://pdfs.semanticscholar.org/facd/bd4b410670431e3f0ec2cf3dabcc7ef55545.pdf. Accessed 27 August 2019.

Goyal, S., & Sharma, V. (2014). Byzantine Attack on Wireless Mesh Networks: a Survey, *International Journal of Science, Engineering and Technology Research*, *3*(12), 3260-3264, December. Available at: http://ijsetr.org/wp-content/uploads/2014/12/IJSETR-VOL-3-ISSUE-12-3260-3264.pdf. Accessed 27 August 2019.

Grandison, T., & Sloman, M. (2000). A Survey of Trust in Internet Applications, *IEEE Communications Surveys & Tutorials*, *3*(4), 2-16, January. DOI: 10.1109/COMST.2000.5340804

Han, K., Ravindran, B., & Jensen, E. D. (2007). Byzantine-Tolerant Point-to-Point Information Propagation in Untrustworthy and Unreliable Networks, September 2007. Available at: https://www.researchgate.net/publication/220909574_Byzantine-Tolerant_Information_Propagation_in_Untrustworthy_and_Unreliable_Networks. Accessed 26 August 2019.

Henderson, H. (2007). Mercury introduces better systems after Muliaga death, *New Zealand Herald*, 2 July. Available at: https://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10447834. Accessed 27 August 2019.

IEC. (1986). IEC60880: Software for computers in the safety systems of nuclear power stations, International Electrotechnical Commission. Available at: https://webstore.iec.ch/publication/18251. Accessed 27 August 2019.

IEC. (2010). IEC61508: Functional safety of electrical/electronic/programmable electronic safety-related systems, International Electrotechnical Commission. Available at: https://www.iec.ch/functionalsafety/. Accessed 27 August 2019.

IEC. (2014). IEC62645: Nuclear power plants - Instrumentation and control systems - Requirements for security programmes for computer-based systems, International Electrotechnical Commission. Available at: https://webstore.iec.ch/publication/7311. Accessed 27 August 2019.

ISO/IEC [International Organization for Standardization/International Electrotechnical Commission]. (no date). ISO 27000 Series of Standards. Available at: https://www.itgovernance.co.uk/iso27000-family. Accessed 26 August 2019.

Jayaswal, B. K., & Patton, P. C. (2006). *Design for Trustworthy Software: Tools, Techniques, and Methodology of Developing Robust Software*. Upper Saddle River, NJ: Prentice Hall. ISBN 978-0131872509.

Katwala, A. (2019). Here's how GCHQ scours Huawei hardware for malicious code, *Wired*, 22 February. Available at: https://www.wired.co.uk/article/huawei-gchq-security-evaluation-uk. Accessed 26 August 2019.

Kourie, D. G., & Watson, B. W. (2012). *Correctness-By-Construction Approach to Programming*. Springer. ISBN 9783642279195.

Letts, S. (2019). China policy on Australian coal is 'as dark and impenetrable as night' and that's how it wants it, *ABC News*, 25 February. Available at: https://www.abc.net.au/news/2019-02-25/china-policy-on-australian-coal-dark-and-impenetrable/10843148. Accessed 23 March 2019.

Lipovsky, R., & Cherepanov, A. (2016). BlackEnergy trojan strikes again: Attacks Ukrainian electric power industry, *welivesecurity*. Available at: (https://www.welivesecurity.com/2016/01/04/blackenergy-trojan-strikes-again-attacks-ukrainian-electric-power-industry. Accessed 26 March 2019.

Marsh, S. P. (1994). *Formalising Trust as a Computational Concept*, PhD Thesis, University of Stirling. Available at: https://www.nr.no/~abie/Papers/TR133.pdf. Accessed 23 March 2019.

Matt, C., Benlian, A., & Hess, T. (2015). Digital Transformation Strategies, *Business & Information Systems Engineering*, *57*(5), 339-343, October. DOI: 10.1007/s12599-015-0401-5

Microsoft. (2019). Get instantaneous behavioral analytics and anomaly detection, 2 April. Available at: https://docs.microsoft.com/en-us/cloud-app-security/anomaly-detection-policy. Accessed 27 August 2019.

Miller, D. (2019). Information Dominance: The Philosophy, *GPF: Global Policy Forum*, 29 December. Available at: https://www.globalpolicy.org/component/content/article/154/26581.html. Accessed 23 March 2019.

Mimoso, M. (2017). Platinum APT First to Abuse Intel Chip Management Feature, *Threatpost News Wrap*, 9 June. Available at: https://threatpost.com/platinum-apt-first-to-abuse-intel-chip-management-feature/126166/. Accessed 23 March 2019.

Mühleisen, M. (2018). The Long and Short of The Digital Revolution, *Finance & Development*, *55*(2), 4-8, June.

National Journal. (2018). The Balkanization of Global Tech, *National Journal*, 30 April. Available at: https://www.nationaljournal.com/s/667253/balkanization-global-tech. Accessed 23 March 2019.

NDIA [National Defense Industrial Association]. (2017). Team 2 Summary: Trustable Access to Leading Edge Technology, *NDIA Trusted Microelectronics Joint Working Group*, July. Available at: https://www.ndia.org/divisions/%20working-groups/tmejwg/final-team-reports. Accessed 23 March 2019.

Ning, S., & Wu, H. (2017). China: Cybersecurity 2017. Available at: https://iclg.com/practice-areas/cybersecurity-laws-and-regulations/china. Accessed 27 August 2019.

NSTC [National Science and Technology Council]. (2011). Trustworthy Cyberspace: Strategic Plan for the Federal Cybersecurity Research and Development Program. Available at: https://www.nitrd.gov/pubs/Fed_Cybersecurity_RD_Strategic_Plan_2011.pdf. Accessed 27 August 2019.

OECD. (2016). The Economic Impact of Local Content Requirements. Available at: https://www.oecd.org/trade/topics/local-content-requirements/. Accessed 27 August 2019.

Otuoze, A., Mustafa, M. W., & Larik, R. M. (2018). Smart grids security challenges: Classification by sources of threats, *Journal of Electrical Systems and Information Technology*, *5*(3), 468–483, December. DOI: 10.1016/j.jesit.2018.01.001

Palo Alto. (2014). Getting Started With a Zero Trust Approach to Network Security, 25 March. Available at: https://www.bankinfosecurity.com/whitepapers/getting-started-zero-trust-approach-to-network-security-w-973. Accessed 27 August 2019.

Park, D., Summers, J., & Walstrom, M. (2017). Cyberattack on Critical Infrastructure: Russia and the Ukrainian Power Grid Attacks, *University of Washington, The Henry M. Jackson School of International Studies*, 11 October. Available at: https://jsis.washington.edu/news/cyberattack-critical-infrastructure-russia-ukrainian-power-grid-attacks/. Accessed 23 March 2019.

Qiang, C. Z.-W., Rossotto, C. M., & Kimura, K. (2009). Economic Impacts of Broadband, *World Bank Report*, Chapter 3. Available at: https://siteresources.worldbank.org/EXTIC4D/Resources/IC4D_Broadband_35_50.pdf. Accessed 23 March 2019.

Saadi, R., Rahaman, M. A., Issarny, V., & Toninelli, A. (2011). Composing Trust Models towards Interoperable Trust Management. In: Wakeman, I., Gudes, E., Jensen, C. D., & Crampton, J. (eds), *Trust management V*. IFIPTM 2011, *IFIP Advances in Information and Communication Technology*, *358*, 51–66. Berlin: Springer. DOI: /10.1007/978-3-642-22200-9_7

Sen, J. (2010). A Distributed Trust Management Framework for Detecting Malicious Packet Dropping Nodes in a Mobile Ad-hoc Network, *International Journal of Network Security & Its Applications*, *2*(4), 92-104, October. DOI: 10.5121/ijnsa.2010.2408

Sherwood, J., Clark, A., & Lynas, D. (2005). *Enterprise Security Architecture: A Business-Driven Approach*, CRC Press. ISBN:9781578203185

Sindhuja, K., Nasrinbanu, A., & Elavarasi, K. (2015). Survey on Malicious Node Detection and Reliable Data Fusion in MANET, *International Journal of Scientific Research Engineering & Technology*, *4*(3), 202-205, March.

Skopik, F., Ma, Z., Bleier, T., & Gruneis, H. (2012). A Survey on Threats and Vulnerabilities in Smart Metering Infrastructures, *International Journal of Smart Grid and Clean Energy*, *1*(1), 22-28.

Speicher, C. (2011). Security Fabric – Tailored Trustworthy Space, Part1: Flexibility Based on Policy Management, *Grid Interop Forum 2011*. Available at: https://www.gridwiseac.org/pdfs/forum_papers11/speicher_paper_part1_gi11.pdf. Accessed 26 March 2019.

Teikari, O., & Nevalainen, R. (2014). Comparison of software safety standards IEC 61508-3 and IEC 62138, *VTT Research Report* VTT-R-03820-14. Available at: https://www.vtt.fi/inf/julkaisut/muut/2014/VTT-R-03820-14.pdf. Accessed 23 March 2019.

US DoD. (1983). Department of Defence Trusted Computer System Evaluation Criteria. Available at: https://csrc.nist.gov/CSRC/media/Publications/white-paper/1985/12/26/dod-rainbow-series/final/documents/std001.txt. Accessed 26 August 2019.

Varghese, S. (2019). Huawei cyber testing centre rejection by Australia 'an old story', *ITWire*, 6 March. Available at: https://www.itwire.com/government-tech-policy/86272-huawei-cyber-testing-centre-offer-to-australia-an-old-story.html. Accessed 26 March 2019.

Volz, D. (2017). Trump signs into law U.S. government ban on Kaspersky Lab software, *Reuters*, 13 December. Available at: https://www.reuters.com/article/us-usa-cyber-kaspersky/trump-signs-into-law-u-s-government-ban-on-kaspersky-lab-software-idUSKBN1E62V4. Accessed 23 March 2019.

Wlodarczak, P. (2017). Cyber Immunity - A Bio-Inspired Cyber Defense System, *Lecture Notes in Computer Science*, *10209*, April. DOI: 10.1007/978-3-319-56154-7_19

Worstall, T. (2013). If Apple Brought iPhone Manufacturing To The US It Would Cost Them $4.2 billion, *Forbes*, 25 September. Available at: https://www.forbes.com/sites/timworstall/2013/09/25/if-apple-brought-iphone-manufacturing-to-the-us-it-would-cost-them-4-2-billion/#1fdce952115f. Accessed 23 March 2019.

Zeadally, S., Pathan, A-S. K., Alcaraz, C., & Badra, M. (2013). Towards Privacy Protection in Smart Grid, *Wireless Personal Communications*, *73*(1), 23-50, November.

Zhang, D. (2017). Intrusion Tolerance for CT Cloud Security, *RSA Conference*, Abu Dhabi.

Zouridaki, C., Mark, B. L., & Hejmo, M. (2007). Byzantine Robust Trust Establishment for Mobile Ad-hoc Networks, *Telecommunications Systems*, *35*(3-4), 189-206, August. DOI: 10.1007/s11235-007-9047-z

Zouridaki, C., Mark, B. L., Hejmo, M., & Thomas, R. K. (2007). HERMES: A Quantitative Trust Establishment Framework for Reliable Data Packet Delivery in MANETs, *Journal of Computer Security*, Special Issue on Security of Ad-Hoc and Sensor Networks, *15*(1), 3-38, January. DOI: 10.3233/JCS-2007-15102