Securing the Future: Integrating Quantum Randomness into Enterprise Cryptographic Systems

Richard Searle[1], Ramy Shelbaya[2], George Dunlop[2]
Fortanix, Inc.[1], Quantum Dice Limited[2]
Oxford, UK

Abstract

Modern hardware security modules (HSMs) rely on approved techniques for random bit generation. Developments in quantum random number generators (QRNGs) now support efficient integration of external entropy sources to provide additional features for HSM customers. In this industry case study, we report various pathways for the integration of a QRNG solution with a certified HSM. We present a preliminary assessment of the benefits contributed by the integration and conclude by identifying areas for future technical evaluation.

1. Introduction

The modern digital economy depends upon the secure storage and transmission of massive volumes of data, using encryption methods predicated on cryptographic keys. To protect the associated cryptographic keys used to encipher raw data, enterprises rely upon the use of hardware security modules (HSMs). NIST defines an HSM as "a physical computing device that safeguards and manages cryptographic keys and provides cryptographic processing," adding that an HSM "is or contains a cryptographic module" [1].

Under the NIST definition, a cryptographic module is deemed to comprise "the set of hardware, software, and/or firmware that implements approved cryptographic functions (including key generation) that are contained within the cryptographic boundary of the module" [1]. Hence, an HSM may consist of either software or hardware conforming with the security requirements established by the related international standards ISO/IEC 19790:2012[1] [2] and ISO/IEC 24759:2017 [3].

The reliability and security of the cryptographic keys that the HSM generates is paramount to the overall security of the systems that it serves to protect. These keys need to be generated via a fully random process and their security is directly tied to the randomness that creates them. The quantity of randomness required for a key will differ from between cryptographic methods and will generally be defined by the relevant standard for the protocol in question. However, with the rapid increase in computational power, the amount of randomness required for each key is constantly increasing, especially for post-quantum encryption algorithms. Additionally, the refresh-rate for these keys, the number of times a new key needs to be created within a given period, is also accelerating.

The process of cryptographic key generation within a cryptographic module requires reference to a source of high-quality entropy to obtain a secret random initialization value, or seed, as the input to a deterministic random bit generator (DRBG). For robust cryptographic security, the initial seed should not be reused, should remain unknown, and should be a random, non-deterministic, sequence of values. The source of noise from which the uncertainty in the output of the entropy source derives is, therefore, the root of security for both the entropy source and the dependent DRBG underpinning the security of the cryptographic module [4].

2. Entropy in key generation

The concept of entropy traces its genesis in the physical domain to Clausius (1864), but it remains a multifarious construct with different manifestations across a broad spectrum of rigorous scientific disciplines [5].

For the purposes of cryptographic module validation under FIPS PUB 140-3 [6], a branch of Rényi entropy, referred to as min-entropy, $H^{(M)}$, is adopted as a conservative measure of "the effectiveness of the strategy of guessing the most likely output of the entropy source" [4]. This min-entropy, is measured in bits and can be mathematically defined as:

$$H_{\infty}^{(M)} = -\log_2\left(\max_i P(X)(x_i)\right) \quad (1)$$

For a particular set of outcomes X, where each outcome can have a probability $p(x_i)$, this can be intuitively understood by considering a fully deterministic process whereby, if a particular outcome, x_j, is certain to happen $\left(p(x_j) = 1\right)$, the

[1] IOS/IEC 19790:2012 is expected to be replaced imminently by draft standard IOS/IEC FDIS 19790.

 | DOI: 10.20533/jitst.2046.3723.2024.0101

min-entropy will then be 0 bits as one would expect. Conversely, the more unpredictable a process is – i.e., the smaller the probability of any particular outcome – the bigger the min-entropy, $H^{(M)}$, gets, such that:

$$\lim_{H^{(M)}\to\infty} p(x_j) = 0 \qquad (2)$$

To appropriately seed a DRBG and produce reliably secure cryptographic keys, the min-entropy generated must be consistent. Maintenance of entropy quality is important since any process of randomness extraction assumes a certain level of min-entropy. DRBGs cannot create entropy and can only use what is made available to them. When given insufficient entropy, cryptographic systems can exhibit significant vulnerability, as described within the literature [7, 8].

2.1. Approved entropy sources for HSMs

The entropy source model implemented by HSM solutions, as defined by NIST [4], is illustrated in Figure 1:

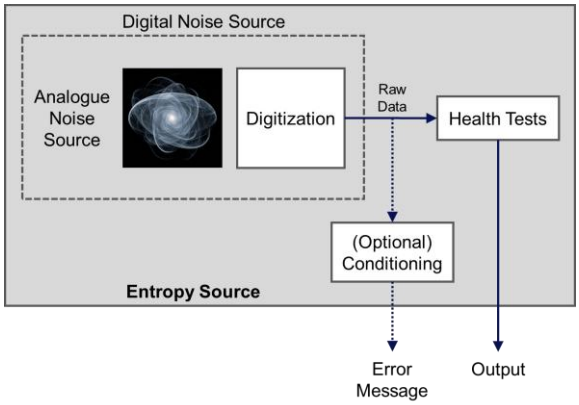

Figure 1. NIST entropy source model for DRBGs

The entropy source model comprises four, discrete, components; a noise source, a digitization component to transform the input noise into bits, an optional conditioning stage to enhance entropy quality post-transformation, and a health test to verify the quality of the entropy available for the cryptographic module.

The output from the entropy source is the string of random values used as an initial seed by the DRBG within the HSM. Error messages generated by the health test component of the entropy source enable rejection of unsatisfactory entropy values to maintain the integrity of dependent cryptographic processes.

2.2. Limitations of classical entropy sources

Generating consistent entropy for DRBG consumption is non-trivial. Various methods have been developed to try to estimate the amount of entropy that a physical process can generate.

The relevant measure here is usually the min-entropy of the output. This is essential because, as the name implies, DRBGs are incapable of adding any entropy, so the security, and unpredictability, of their outputs are intimately tied Although there exist a number of relevant standards that have been established for the analysis and assessment of the output of entropy sources, such as the SP800-90B recommendation maintained by NIST [4], there are still a number of limitations that need to be taken into consideration when choosing the appropriate entropy sources for HSM devices.

For classical entropy sources, the only analysis that can be done on their output is a statistical one, where there is either a statistical model, incorporating several assumptions regarding how the source is supposed to operate, or the user must undertake statistical tests on the output of the source to estimate its properties. This analysis must be done *a priori*, in the sense that it is done once, at the point at which the entropy source is tested during manufacture. It is then assumed that the source will continue to perform with the same output capability in all the different conditions that it may experience during its operational lifetime. However, it is well known that hardware systems are prone to several variabilities both at their time of manufacturing and over their lifetime. This variability in performance stems from a combination of specific environmental effects and general degradation of components due to ageing. It is also possible to maliciously manipulate the hardware components of the entropy source, either physically or remotely, to induce a deviation in its expected behavior.

Existing health tests, that are in principle designed to detect the above deviations, are also dependent upon statistical tests applied on the output of the source to determine whether the source is operating as designed. Here, again, the issue is that these tests are retrospective – i.e., they can only say anything meaningful about the historical output, not about what is going to be produced by the source in future. Furthermore, such tests are also bound by the general limitation of any statistical analysis in terms of their effective validity and reliability.

2.3. The potential of quantum mechanics

The limitations and vulnerabilities are not purely theoretical or academic in nature. Entropy sources, which are supposedly compliant with the relevant standards for entropy source validation (ESV) have failed in real-world security systems and devices, producing less entropy than required and compromising the quality and security of the cryptographic primitives that they are being used to generate. Consequently, the integrity of the

 | DOI: 10.20533/jitst.2046.3723.2024.0101

cryptographic systems which depend on them can be undermined, solely, by the performance of the entropy source used to seed cryptographic key generation processes.

To that end, the search for a better entropy source is continuous, and relevant industry standards are constantly being revised and updated (albeit with insufficient speed). An interesting candidate for more robust entropy sources can be found in the incredible properties characterizing quantum mechanical systems. In the quantum domain, randomness is built into the physical laws that govern how these systems operate and interact. For example, there is no way to predict when a radioactive nucleus will emit a burst of radiation, it is a fundamentally unknowable property of the universe and is thus a candidate for a system required to provide true randomness. This is where the technology of quantum random number generators (QRNGs) [9] demonstrates potential advantages over established entropy sources in their independence from extraneous influences.

3. Pathways to QRNG integration

To take advantage of the benefits offered by QRNG integration, efficient pathways for secure entropy consumption by HSMs must exist. These pathways can be established via software-based interaction between the HSM and the QRNG, or via integration of QRNG hardware within a secure enclosure, in accordance with FIPS 140-3 criteria [6] for physical HSM appliances.

3.1. Software

Acquisition of initial seed values from an independent QRNG entropy source requires support for standard HSM interface protocols that facilitate efficient integration and familiarity for the HSM user. The PKCS#11 standard provides a platform-independent application programming interface (API) for cryptographic interaction with HSMs.

Modern HSM systems, as featured within the reported case study, typically support RESTful APIs to enable programmatic interaction between software applications and an HSM responsible for cryptographic key generation and security. Where QRNG devices support PKCS#11 compatible API requests, entropy can be obtained remotely using a `getentropy` request [4] via a secure communications channel.

3.2. Hardware

In selecting an appropriate quantum mechanical system as an entropy source, direct integration of radioactive isotopes proves impractical due to a variety of potential risk factors in manufacturing and deployment of an HSM appliance. Hence, to be able to have a practicable QRNG, we need easier systems to integrate and control throughout the lifecycle of an HSM.

Thankfully, the rapid development of hardware allowing the implementation of quantum systems, especially in the field of quantum photonics, has led to the development of easy-to-integrate QRNG systems. For example, Quantum Dice's VERTEX QRNG comes in a standardized half-length PCIe card form factor, which allows for a straightforward hardware integration within any HSM system supporting a spare PCIe slot in its hardware design.

In addition to the fundamental security advantages arising from the way its randomness is generated, the VERTEX system has other practical features including a fully post-processed bitrate of 2.6 Gbps, meaning that it can satisfy existing and future entropy requirements, even as standard recommendations on the size of encryption keys and their refresh rates evolve. The VERTEX PCIe device can achieve this performance with a power consumption of only 9W thanks to its optimized photonic implementation. Additionally, having a PCIe card form factor means that the integration of this QRNG in any hardware security solution requires minimal engineering on the part of the HSM manufacturer.

4. Source-device independence

In principle, quantum processes are inherently truly random, like any other hardware implementation, there is no such thing as an ideal quantum system – there will always be an element of classical noise and fluctuations in the way the system operates. Returning once again to the radioactive decay example, the process itself is heavily dependent on the exact chemical composition of the material. Additionally, detecting the decay itself will have to involve using other hardware components with their own internal noise and fluctuating accuracy. All of this means that an entropy measure based on a particular model for the quantum system is not guaranteed to persist in operation, especially if there are environmental fluctuations. Hence, dependance of a random number generator on a quantum process is not sufficient to guarantee dependable and consistent levels of entropy.

For this reason, various concepts of "device-independence" have been developed which aim to address this challenge by leveraging certain aspects of suitable quantum architectures to mitigate the impact of external disturbances on entropy source outputs. What device-independence means is that fewer assumptions are required about the correct, or reliable functioning of the individual components constituting the quantum architecture. This greatly reduces the possible attack vectors, since it means that if some of the components were to be compromised, they would

 | DOI: 10.20533/jitst.2046.3723.2024.0101

not impact the amount of entropy that the system can be expected to produce. For example, one particularly sensitive part of any quantum system is the source of the quantum state itself. The hardware used to establish the quantum state comprises especially delicate components that can be very sensitive to even slight disturbances. A compromised generation of the quantum state will inevitably have a significant effect on the quality of the entropy that the system can generate.

To mitigate this concern, a source-device independent system eliminates the need to rely on the performance of the device used to generate the quantum state. The exact methods by which independence can be achieved are myriad and depend on the specific system under consideration. A common approach is to leverage the unique characteristics of quantum physics that allow you to gain some information about a quantum state without fulling "collapsing" the randomness inherent in its final measurement.

4.1. Continuous entropy estimation

Quantum Dice's patented source-device independent self-certification (DISC) protocol which enables all its commercial QRNG offerings, including the VERTEX product line, allows for an instant continuous entropy estimation – based on a source-device independent approach [10]. If the process generating the quantum state fluctuates through classical noise, this is detected and taken into consideration. If we compare this to the health tests, or entropy checks that are performed classically within contemporary cryptographic systems; the DISC protocol allows for instantaneous real-time verification on the data that is currently being generated, instead of simply performing a statistical check on the historical output of the device. This allows users and systems that rely on DISC-backed QRNGs to have improved visibility of the performance, and therefore integrity, of their key generation devices and to detect issues immediately as they occur. Consequently, DISC-backed QRNGs produce two outputs: 1) entropy data derived from measurement of the quantum random process, and 2) certification data that allows the end-user to have a continuous measure of the entropy that is being generated by the source.

5. Case Study

As an example of the potential application of QRNG technology, this case study considers the integration of Quantum Dice QRNG solutions with Fortanix Data Security Manager™ (Fortanix DSM). Fortanix DSM is a commercial HSM system available in a variety of form factors, including a physical appliance conforming to secure hardware criteria established under FIPS 140-3 Level 3 [6], global software-as-a-service (SaaS), and supported software-based deployment using virtual machine instances. Different instantiations of Fortanix DSM are underpinned by common product functionality to support enterprise cryptography at scale. For cryptographic key generation, all forms rely on a DRBG that is supplied with random seed values by a designated FIPS-approved entropy source.

In this case study we report separate methods for integration of a Quantum Dice QRNG entropy source with Fortanix DSM to provide two desirable product features for HSM customers: (1) The ability for HSM customers to select a source-device independent QRNG as the preferred entropy source for DRBG seeding in cryptographic key generation processes. (2) Use of the HSM as a server node for secure entropy-as-a-service (EaaS) provision to consumer applications, via a secure API that passes initial seed values from the QRNG to external applications or devices.

5.1. Reported system architecture

The case study system architecture comprises an instance of Fortanix DSM that operates, securely, within memory using the Intel® Software Guard Extensions (Intel® SGX) technology to create trusted execution environments (TEEs). The standard Fortanix DSM implementation utilizes a FIPS-approved true random number generator (TRNG), however, a Quantum Dice QRNG can be securely connected to Fortanix DSM to function as an alternative entropy source.

One option is to integrate the VERTEX PCIe card inside the Fortanix FX3400-series appliance incorporating FIPS 140-3 Level 3 standard security.

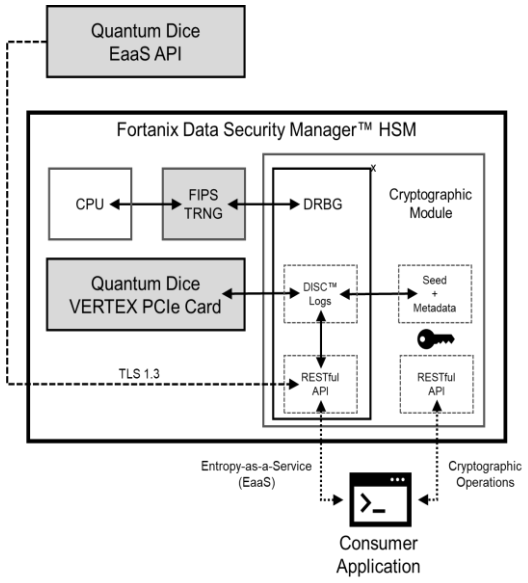

Figure 2. Case study system architecture

A second option is to establish remote EaaS connection between a QRNG source device and the HSM. Here, Fortanix DSM can consume external entropy via a RESTful API supported by a PKCS#11 cryptographic library. This integration pathway yields the potential for multiple, differentiated, QRNG sources to be available for consumption by the HSM. The reported case study system architecture is represented in Figure 2:

5.2. Entropy validation

The case study integration takes advantage of the DISC protocol by incorporating entropy certification as metadata within custom key attributes during the key generation process. Where new keys are created via API requests to the HSM, or the HSM user interface, the quality of the applied entropy can be verified to maintain a consistent security posture.

In addition to providing a unique signature that allows the end-use application to assure the suitability of each key that is being generated, the certification signal also allows the system to monitor the performance of the quantum source. While not impacting the quality of the final output, fluctuation of the certification signal indicates that there may be some external disturbance affecting the entropy source hardware.

6. User Benefits of QRNG Integration

The Fortanix DSM HSM incorporates a FIPS-certified TRNG as the default system entropy source, the ability for customers to select an alternative QRNG source for initial seed value generation provides potential benefits that require further evaluation in the field.

6.1. Configurable entropy source selection

Since the QRNG is based upon a physical quantum system, with inherent uncertainty regarding its temporal state, live verification of the quality of the entropy which is then used by the DRBG mitigates potential issues derived from insufficient min-entropy in key generation. Enabling HSM customers to select from one or more QRNG source providers, whether on-board the HSM through PCIe card integration, or from a FIPS secured EaaS node, allows for control, monitoring, and subsequent audit of entropy inputs via the HSM log data.

The ability to harvest entropy from multiple quantum sources, with quality guarantees bound to the contingent cryptographic secrets, provides the ability for an HSM to derive fully quantum secure cryptographic keys, consistent with a planned migration to post-quantum cryptography in accordance with NIST SP 1800-38A draft guidelines.

6.2. Source EaaS distribution

EaaS is an emerging framework for distributed delivery of high-quality randomness to end-point devices, such as internet-of-things (IoT) endpoints or low-resource systems, which either struggle to reliably generate their own entropy or cannot host an on-board entropy source. EaaS provides developers a way to guarantee the quality of cryptography under context-dependent platform constraints.

Integrating a DISC-backed QRNG with a secure HSM to allow the operation of a secure EaaS will be transformational for connected applications as it combines the best of strengths of each product to provide a secure, scalable, and verifiable source of entropy. Examples of EaaS enabled use cases include remote on-device key generation within IoT manufacturing, secure entropy acquisition within Monte-Carlo simulation applications in financial services, and prevention of entropy starvation within high-utilization virtualized computing resources.

7. Conclusions

Modern HSM architectures, with rich interface support, as demonstrated by the case study example of Fortanix DSM, coupled with PKSC#11 compatibility of miniaturized QRNG devices, such as the Quantum Dice VERTEX PCIe card, offer potentially beneficial feature enhancements for users.

HSM integration of QRNG-derived entropy with verifiable quality guarantees enhances compliance monitoring by linking source entropy to individual cryptographic keys using custom metadata attributes. Thus, individual keys can be evaluated by consuming applications prior to use. Furthermore, EaaS consumer applications and devices can reject compromised entropy inputs to maintain min-entropy in cryptographic key generation and other dependent processes.

In addition to the integration pathways discussed by this case study, there exists future opportunity to evaluate the benefits of implementing secure, certified, QRNG entropy within HSM systems for use in artificial intelligence (AI) applications and data privacy settings. Comparison of information-theoretic entropy in datasets, and introduction of noise to maintain differential privacy norms, may prove to be novel applications for QRNG entropy sources, secured via trusted HSMs.

8. References

[1] Barker, E. and Barker, W. C. (2019). Recommendation for Key Management: Part 2 – Best Practices for Key Management Organizations. NIST Special Publication 800-57 Part 2: Revision 1, National Institute of Standards and Technology (NIST).

 | DOI: 10.20533/jitst.2046.3723.2024.0101

[2] International Organisation for Standardization (2015). Information technology – Security techniques – Security requirements for cryptographic modules. ISO/IEC 19790: 2012(2).

[3] International Organisation for Standardization (2017). Information Technology – Security Techniques – Test requirements for cryptographic modules. ISO/IEC 24759:2017(3).

[4] Turan, M. S., Barker, E., Kelsey, J., McKay, K. A., Baish, M. L. and Boyle, M. (2018). Recommendation for the Entropy Sources Used for Random Bit Generation. NIST Special Publication 800-90B, National Institute of Standards and Technology (NIST).

[5] Zolfraghari, B., Bibak, K. and Koshiba, T. (2022). The Odyssey of Entropy: Cryptography. *Entropy*, 24(2), 266.

[6] National Institute of Standards and Technology (2019). Security Requirements for Cryptographic Modules. (Federal Information Processing Standard) FIPS PUB 140-3, National Institute of Standards and Technology (NIST).

[7] Shikata, J. (2017). Tighter Bounds on Entropy of Secret Keys in Authentication Codes. 2017 IEEE Information Theory Workshop (ITW), Kaohsiung, pp. 259-263.

[8] Benadjila, R. and Ebalard, A. (2023). Randomness of random in Cisco ASA. Unpublished (Cryptology ePrint Archive), Paper 2023/912.

[9] Jacak, M. M., Jóźwiak, P., Niemczuk, J. and Jacak, J. E. (2021). Quantum generators of random numbers. *Scientific Reports*, 11(16108).

[10] Drahi, D., Walk, N., Hoban, M. J., Fedorov, A. K., Shakhovoy, R., Feimov, A., Kurochkin, Y., Kolthammer, W. S., Nunn, J., Barrett, J., and Walmsley, I. A. (2020). Certified Quantum Random Numbers from Untrusted Light. *Physical Review X*, 10(041048).

 | DOI: 10.20533/jitst.2046.3723.2024.0101