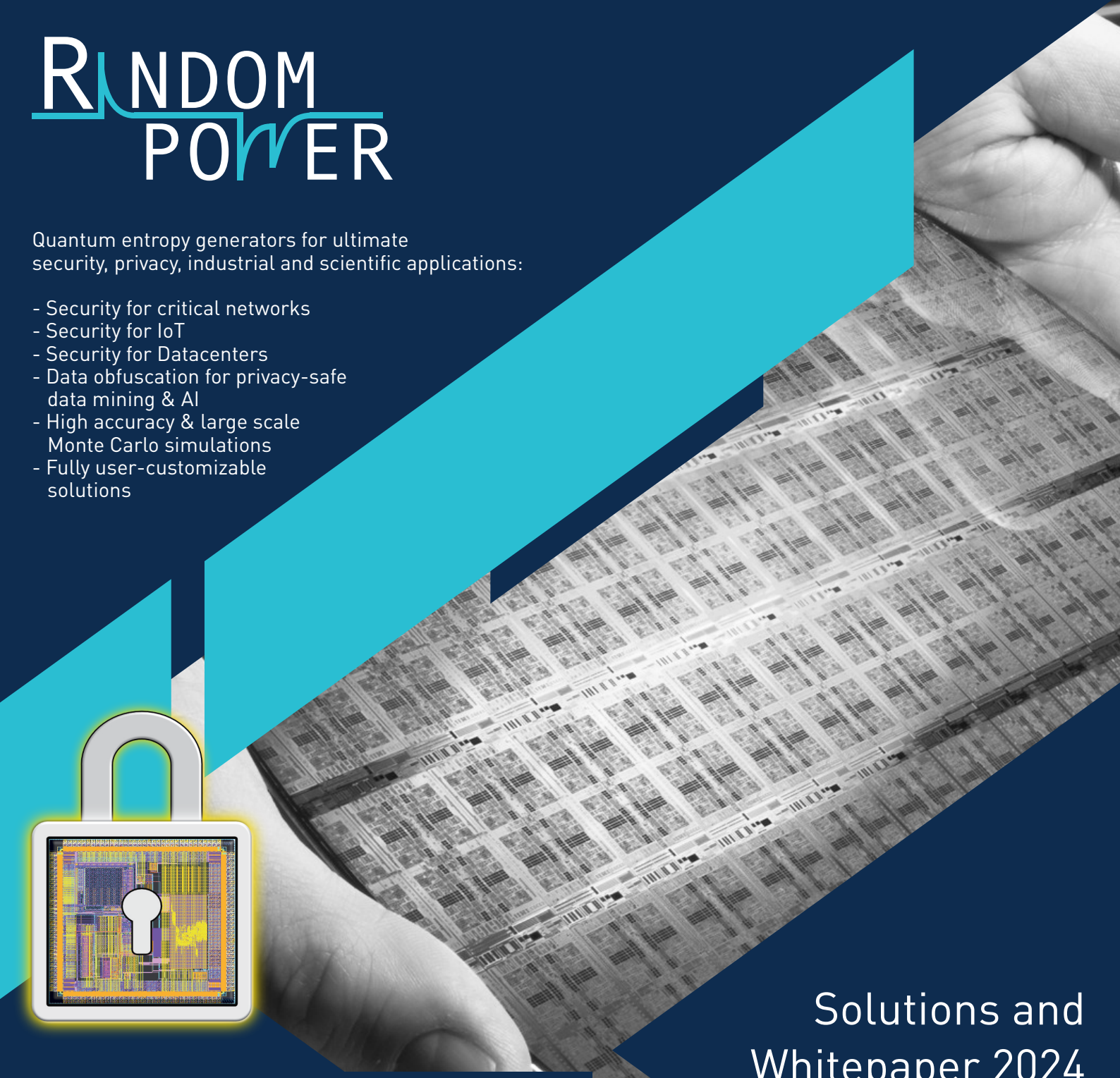
RANDOM POWER

Quantum entropy generators for ultimate security, privacy, industrial and scientific applications:

- Security for critical networks
- Security for IoT
- Security for Datacenters
- Data obfuscation for privacy-safe data mining & AI
- High accuracy & large scale Monte Carlo simulations
- Fully user-customizable solutions



Solutions and Whitepaper 2024

Secure your World with a Quantum True RNG

ATTRACT · AGH · UNIVERSITÀ DEGLI STUDI DI MILANO · imasenic · RANDOM POWER

SECO · E4 COMPUTER ENGINEERING · weeroc · NAGRA KUDELSKI GROUP · FONDAZIONE BRUNO KESSLER

About us

Origins

Random Power springs off a long standing experience in the development of Silicon Detectors for Sub-atomic particles, notably for experiments at CERN, the European Laboratory for Particle Physics. Know-how and technology developed for blue sky research was turned into interdisciplinary and industrial applications, thanks to the support of a series of grants by the European Commission (EC), in an exciting loop. As a final step, with an act of pure serendipity, the principle at the base of Random Power emerged by a flash of ingenuity.

2016

The principle at the base of RandomPower emerges, as result of a genuine serendipity event. Initial tests performed with lab equipment

2017

Unpredictability of the generated random bit sequence is assessed using the NIST standards

- A **demo board** is designed, commissioned and qualified;
- Italian patent filing completed (October)
- launch at the CyberSecurity week in Le Hague (October)
- **submission** of the **ATTRACT Phase 1** proposal (October) (100 kEUR)

2018

- **Approval** & **kick-off** of the **ATTRACT Phase 1** proposal (May)
- design, commissioning and production of the single generator board;
- **End** of the **ATTRACT Phase 1** project (October)
- **Full characterisation** of the single generator board
- **winner** of the **Start-Cup competition** (regional level; 20 kEUR)
- **winner** of two special prizes by investors at PNI, start-up competition at **national level**
- **Italian patent granted**

2019

- **approval** of the **ATTRACT Phase 2** proposal (January) (**2 MEUR**)
- **kick-off** of the ATTRACT Phase 2 (May)
- proto-farm commissioned
- implementation of the TESTU01 suite, complementing the NIST test (1.5 Tb qualified)
- **winner** of the Falling Walls venture int'l competition (November)
- **US and EU patent granted**

2020

- **Chip delivery (Q2)**
- Chip qualification and packaging (Q2-Q3)
- 64x board product grade qualification (Q2-Q3)
- v1.0 of the FIPS compliant sw architecture for the multi-age board
- **FIPS 140-3 certification started**
- **Business plan v2.0**
- **End of the ATTRACT Phase 2 (fall)**
- **Execution of the next investment round**
- **Engineer co-development of applications**
- **GO TO MARKET!**

2021

- first implementation of real-time sanity checks
- **company establishment** (June)
- **investment (200 kEUR)** by **LifTT**, our VC (June)
- **Submission** of the **ATTRACT Phase 2** proposal (September)

2022

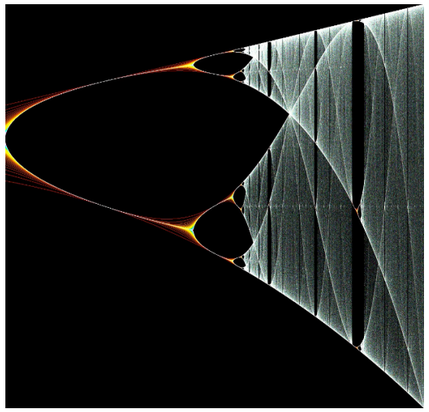
- real-time sanity check implemented
- design of the multi-gen board completed (April) v1.0
- commissioning of v1.0 started (July) ; qualification expected by December
- **chip submission in mid-September**
- **Engineering of the next investment round (including Exploitation agreements)**

2023

2024

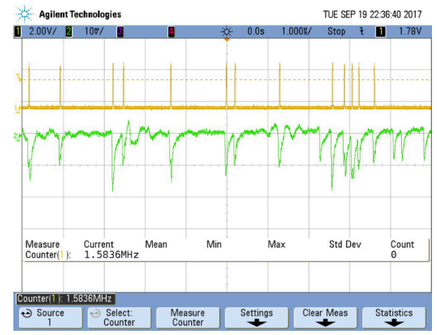
A "Quantum Coin Flipper" for privacy, security, AI and data science

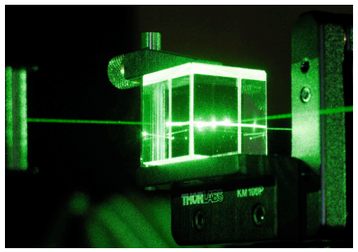
Nowadays randomness is one of the fundamental ingredients at the core of the information technology, powering all the encryption and authentication algorithms that protect our digital lives. Furthermore, novel technologies are being developed for ensuring privacy and confidentiality of the big data used for training complex AI models, and all of them requires high-quality randomness. Finally, the modelization of complex systems in finance, industry and science is often based on "Monte Carlo" simulations, whose accuracy depends on the quality of the required Randomness. In order to guarantee such quality, every computer uses various techniques for grabbing randomness from user-related, network-related or electronics-noise caused events, but there is only one case in which the entropy is really guarantee by the fundamental laws of nature and that is the quantum processes. By extracting endless sequences of bits from quantum events, such as the quantum tunneling through an energy barrier, the Random Power technology provides top-quality entropy ,that ensures that every encryption algorithms will really perform with the highest level of security allowed by its mathematical fundations and technical implementation.

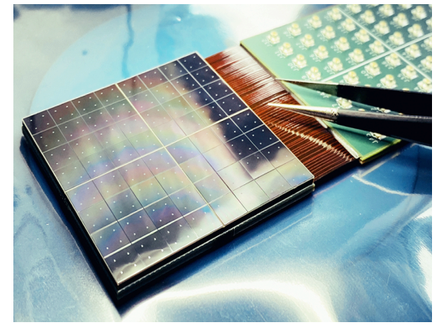


Taking quantum entropy from labs to the market

Being this goal so ambitious and impactful on our techology, Random Power is not alone and several competitors are trying to design devices exploiting quantum processes for entropy generation. They include radioactive decay of unstable nuclei or, more often, photonic technologies. But they are complex, expensive, possibly weak and not compliant with standard CMOS technology, making their miniaturisation and integration in a chip featuring advanced functionalities quite hard if not impossible. Here comes Random Power, that starts as an academic spin-off of research teams from the University of Insubria in Como, Italy and the AGH University of Science and Techology in Krakow, Poland, with top-level experience in the design of CMOS sensors for high-energy physics experiments at CERN. This expertise lead to the design of a technology that exploits the quantum tunneling effect occuring inside semiconductors, thus leading to a design that is real-quantum, but at the same time, it can be easily integrated with the commercial processes for the production of integrated circuits. The result is a line of products that are less expensive and can be integrated inside any commercial PCB just as any other chip, or inside a server as any other PCIe card.







Product Specifications

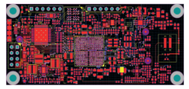
STATE OF THE PROJECT

The platform consists of three elements:

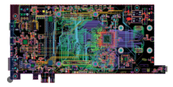
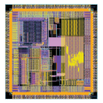
A **single generator board**, with an initial development in an early stage of the project and by now brought to completion, demonstrated the viability of the Random Power concept in an operational environment, complying the U.S. National Institute of Standard and Technology's test suites. The single generator board provides the ideal test instruments.

A **64 generator board**, with a front-end transitioning from discrete to integrated electronics and control through a System-On-Chip, is a customized computer on a board. This system, also because the double I/O through Ethernet or PCI-Express, is a scalable device for Data Centers providing high bit stream rate. The board is FIPS-140-3 compliant by design.

An **ASIC** (Application Specific Integrated Circuit) implementation, providing a low power solution for devices at the edge and Industrial IoT Gateways. The ASIC is FIPS-140-3 compliant by design. It implements an array of generators, the Time-To-Digital converters for time stamping the pulses, bit generation and bit processing through the NIST DRBG procedure (SP800-90 A,B,C). The bit stream can be encrypted (AES-256 protocol).

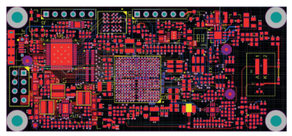




	Single Generator Board	64x Generator Board	ASIC
Dimensions [cm²]	8x3.5	11.1x31.2x2.0	1x1
No. generators	1 array	64 arrays	1 array
Raw bit stream:	100 kbps	32 Mbps	1 Mbps
NIST DRBG output (SP800-90 A,B,C)	NA	1 Gbps	32 Mbps
Control:	Xilinx Spartan 7	Xilinx KRIA K26 SOM	SPI at 24 MHz
I/O:	USB or bits-on-pin	Eth or PCI-Express	SPI at 24 MHz
Power supply:	through the USB (5V, 0.5A)	12V, 8A	5V, 1.8V
Power consumption:	<2.5W	15 W	100 mW
Encryption of the bit stream:	No	Yes (AES-256)	Yes (AES-256)
Specific Features:	• Firmware implemented Real-Time sanity checks (MONOBIT and RUNS) • Auxiliary post-processing through a SHA256 function	• Firmware implemented Real-Time sanity checks (MONOBIT, RUNS, Adaptive proportion test, Repetition Count Test) • Auxiliary post-processing through a SHA256 function • Interface through the Trusted Execution Environment • Temperature control through a Peltier cooler • FIPS-140-3 compliant by design • Possibility to run any user-defined, entropy consumer, application on the embedded Linux OS in both stand-alone and PCI-Express mode	• On Silicon implementation of the NIST Real-Time sanity checks (Adaptive Proportion Test and Repetition Count Test) • On Silicon implementation of the NIST DRBG protocol • Package: QFN100 • FIPS 140-3 compliancy by design; • CAVP (Cryptographic Algorithm Validation Program) completed
State of development:	• Completed • Full qualification of 2 Tb through the NIST and TESTU01 protocols • Single board control through a GUI or mini-farm control implementing also the NIST DRBG procedure (SP800-90 A,B,C)	• Prototype under test • Product grade design expected by June 2024	• Design Completed • Production on going • Delivery expected by April 2024

RaP Single Generator Board




The first implementation of the innovative RandomPower concept. This USB board, smaller than a credit card, can provide up to 100 kHz of quantum entropy with on-board health tests and SHA-256 real-time entropy whitening. The device is compatible with Windows, Linux and MacOS and comes with a testing GUI and a Python SDK. The board is also capable of streaming the entropy to other boards.

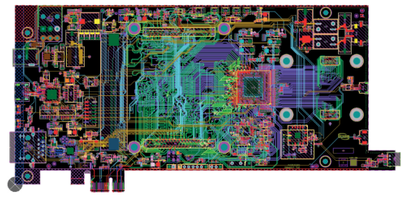
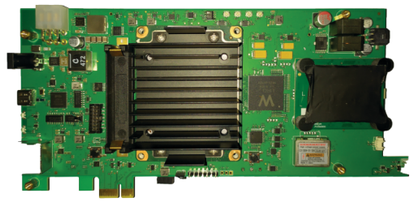
Thanks to the plain-Python code provided in bundle, the board is capable of performing entropy injection into the /dev/random entropy pool in Linux and can provide entropy to the backend of the widespread numpy library for Python. The quality of the randomness has been extensively qualified with the NIST standard test suites and the TESTU01 suite by Pierre l'Ecuyer.

USB 2.0 Interface (FTD2XX serial driver)

Main Features:
- 100 kHz quantum entropy generation
- Embedded Real-Time Sanity Checks
- Auxiliary post-processing SHA256

Applications:
- Linux entropy pool feeding
- Monte Carlo simulations seeding
- Scientific Applications

RaP 64x Generator Board

In the future of data-mining, advanced AI training and cloud computing the privacy of users and corporate data is going to be the main goal. In this landscape, the next-generation algorithms for differential privacy, post-quantum encryption, authentication, and synthetic dataset generation have all in common the request for high-throughput and high-quality unpredictable randomness. Also in the field of high-performance computing the use of stochastic Monte-Carlo methods is widespread for tackling complex problems in finance, industry, big data and science.

The 64x board brings the RandomPower entropy to the datacenter with improved security, reliability and advanced features.

The 64x board, designed to be FIPS 140-3 compliant, embeds a System On Module including a quad-core CPU running Linux, and the in-silicon logic for the quantum entropy generation, starting from the signals that comes from the 64 solid-state sources. The board can work as a stand-alone computer providing entropy and advanced, user-customizable, functionalities to many devices on the LAN or on the Internet. The user can take advantage of ready-to-go applications or can deploy custom software running on Linux. The quantum entropy is secured using hardware-based encryption and stremed to a local or remote process using AES-256 end-to-end encryption; the device control is higly secured using the ARM TrustZone technology.

The board can work with the same level of performance and flexibility inside a server thanks to the compliancy with the PCIe (x1) standard. In this situation the on-board CPU can cooperate with the main server CPU, off-loading randomness-related tasks.

Fast to deploy and easy yo use
- No driver required: both in stand-alone and PCIe mode, the board behaves like a computer connected to your machine via Gigabit Ethernet: application deployment via SSH, data transfer using network protocols and embedded web servers are all supported.

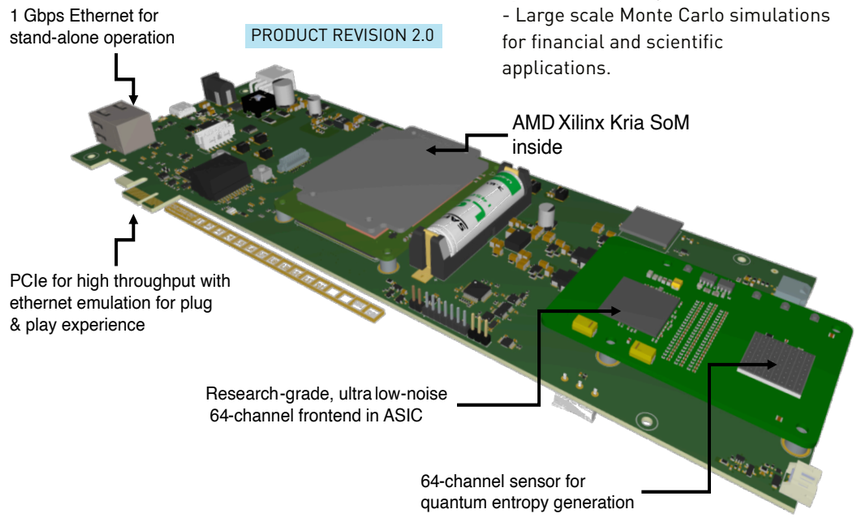
- Ready to use applications: randomness injection into the server entropy pool, integration with numpy random, unified management for multi-board environment, advanced application use cases like differential privacy, synthetic data generation, OpenSSL integration, Quantum Randomness-As-A-Service on LAN, VPN endpoint etc.

- Deploy any application supported by ARM 64 bit: C/C++ compilers, Python 2/3.x, Jupyter Notebook server, Visual Studio Code over SSH, ARM TrustZone Trusted Applications, APT packages etc.

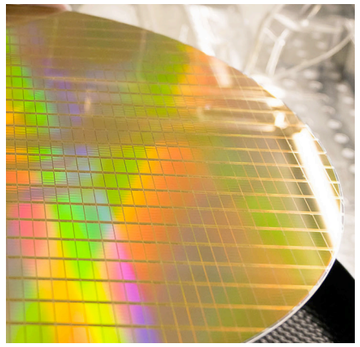
- Ethernet Interface (1 Gbps) on RJ45,
- USB 2.0 console (FTD2XX driver),
- PCI-Express x1 Ethernet Card emulation (Plug&Play on Windows and Linux)

Main Features:
- 64 independent generators with fully automated thermal control
- 32 Mbps quantum entropy generation (1 Gbps from NIST DRBG)
- On-board computer (64 bit ARMv8 quad core, 4 GB RAM, 16 GB eMMC + 1x Real Time Processing Unit) for hosting user-defined workloads
- Advanced on-board security: TPM 2.0, AES, SHA, RSA accelerators, ARM TrustZone extending to the FPGA via secure-AXI, real-time randomness encryption, secure boot, ECDSA-signed EEPROMs, physical security with dedicated MCU and anti-tampering.

Applications:
- The ideal platform for deploying top-level security cloud storage systems and for data obfuscation in large-scale data mining and AI (deep learning, generative AI, AI on medical datasets)
- Differential Privacy, ultra-low-bias synthetic dataset generation
- Large scale Monte Carlo simulations for financial and scientific applications.

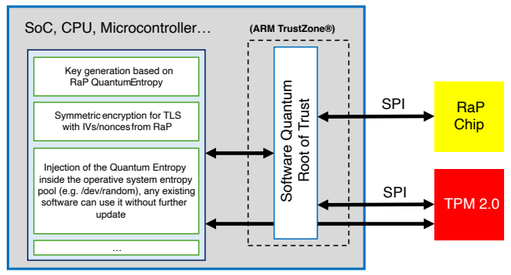


RaP Chip



The RandomPower chip is the flagship of the Random Power technology. It is based on a commercial mixed-signal CMOS process and it can be integrated with the same easiness of any QFN100 integrated circuit. The chip needs a simple single-rail power supply and offers two separate SPI (24 MHz) interfaces, being compatibile with almost any SoC, CPU, FPGA or MCU on the market. The first SPI channel is used for configuration and for setting the AES-256 key used for the random bits encryption, while the second channel is used for the actual streaming of the data. The first channel is used to control several security eFuses and it can be permanently disabled after programming.

The Random Power chip embeds the quantum entropy generation sensor matrix, made of digitally-controlled, independent and indexed single cells, a sub-nanosecond level time-to-digital converter with digital electronics performing entropy 4 bit word extraction, followed by digital processing IP cores including a NIST-compliant certified DRBG (deterministic random bit generator) seeded with the quantum entropy and an AES-256 encryption core that ensures confidentiality of the channel between the host system and the RaP chip. The chip includes digital hardware-implemented health tests that constantly monitor the quality of the output entropy. Integrated with a commercial TPM providing secure key-storage, the RaP chip constitutes the ideal security hardware platform for industrial gateways, critical network devices and IoT systems requiring top-quality locally-generated encryption keys and other security features.



Network devices: VPN endpoints, firewalls, mobile devices and high-end radios

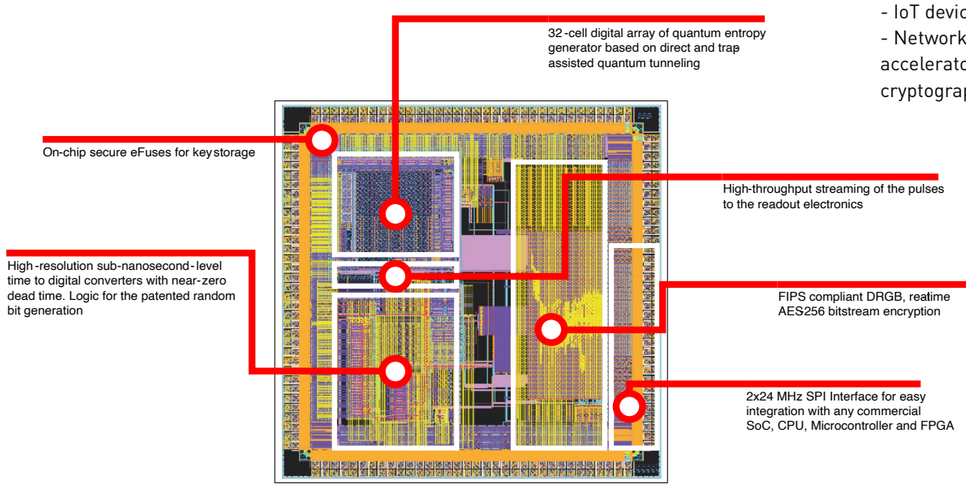
Industry & critical infrastructures: PLC gateways, remote ctrl units

IoT & Professional IT: Smart devices, automotive, smart cities, security cameras

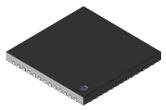
2x SPI 24 MHz

Main Features:
- 1 Mbps on-chip quantum entropy generation, up to 32 Mbps from NIST DRBG
- Embedded AES-256 NIST-compliant Deterministic Random Bits Generator seeded from the quantum entropy source (quantum raw stream available)
- Embedded crypto IP core for real-time encryption of the output bitstream
- 2x24 MHz SPI interface, for ensuring compatibility with any MCU, CPU, SoM, industrial device
- User-defined AES key, security eFuses with possibility to disable the management SPI interface.
- < 100 mW power consumption, 0 - 0 °C temperature range, single rail power supply @ 2.5 V
- Availability of evaluation boards, drivers for many systems, dedicaded support

Applications:
- Used together with a commercial secure element (TPM), allows easy embedding of Quantum Root of Trust in your product, for top-level security
- Industrial gateways, critical infrastructure communication nodes
- IoT devices, smart city, automotive
- Networking: VPN endpoints, TLS accelerators, post-quantum-cryptography security modules



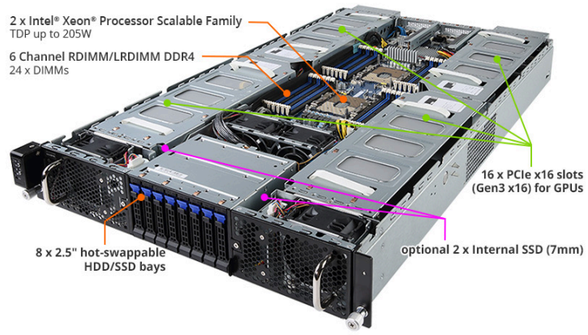
Several patents pending



QFN-100 Package

- PCB-level integration
- SMT module with other crypto-chips…
- Adding SPI to USB/PCIe converter: USB-stick, miniPCIe card, M2 card…

RaP Solutions for Datacenters



E4 COMPUTER ENGINEERING

Thanks to our partner company E4 Computer Engineering, leading the field of high performance computing and datacenter technologies, the RandomPower consortium can provide custom datacenter solutions embedding our Quantum Entropy sources, such as the 64x Multi Generator Board. Our datacenter solutions achieve high density: it is possibile to configure high performance dual CPU servers with up to 8 PCIe board, reaching the number of 512 generators (8 Gbit/s at the DRBG output) in 2 standard rack units. The server solutions are based on the platforms Gigabyte 293 or the SuperMicro Hyper A+ AS-2025, but our boards can be easily integrated also inside other solution, depending on the specific use case.

The servers can be configured in different flavours: optimized for computing, optimized for storage, optimized for network bandwidth (including optical NICs up to 40 Gbit/s), thus satisfying different applications of quantum randomness ranging from cloud storage encryption, differencial-privacy-protected database hosting, privacy applied to AI models and datasets and large-scale Monte Carlo simulations for complex system modelization in science, industry and finance.

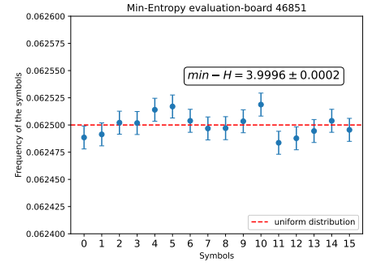
Random bitstream validation

The bit stream produced by the RaP! platform is assessed off-line through a series of test suites and qualified on-line:

- **Off-line qualification** proceeds on the base of the NIST test suite as specified in the SP-800-22 publication, complemented by the Crush, Alphabit and Rabbit suites included in the TESTU01 procedure by Pierre l'Ecuyer and Richard Simard ("TestU01: AC library for empirical testing of random number generators." ACM Transactions on Mathematical Software (TOMS) 33.4 (2007): 1-40.)). The Crush suite was customized by the RaP! team to run over the full bit sequence of a stream, split in sub-samples of 40 GB, defined by the test of the suite requiring the largest sample size.

- **On-line qualification** is based on the Health Tests specified in the NIST documents SP 800-90B (Recommendation for the Entropy Sources Used for Random Bit Generation), namely the Repetition Count and Adaptive Proportion Tests, implanted in firmware not to affect the entropy production rate and applied to strings of 2048 used to seed the Deterministic Random Bit Generation (DRBG) procedure, implemented as of the NIST specification detailed in the SP 800-90B document (Recommendation for Random Bit Generator Construction). Moreover, identification of catastrophic failures based on a RaP! specific implementation of the MONOBIT and RUNS tests are implemented in the single and multiple generator boards.

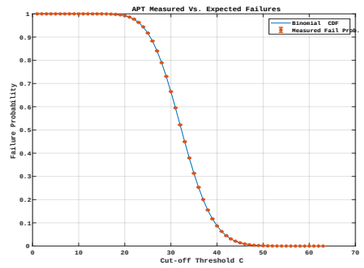
At last, it is worth mentioning that bit samples are processed through the NIST procedure (SP 800-90B) for assessing the bits to be Independent and Identically Distributed. <u>Following a positive outcome, entropy of the raw bits was measured to be in excess of 0.9995 bits/generated bit, averaged over a number of produced devices.</u>



This figure shows the distribution of the frequencies pi of the symbols in a 262.14 MB file generated by one of RaP's boards. The maximum among these values is used to evaluate the min-Entropy according to the formula

$$H - min = -\log_2 \max_{1 \le i \le k} p_i$$

where k is the size of the alphabet used. Perfect unpredictability is associated to H-min = 4 and the result for this board is H-min = 3.9996 ± 0.0002.



The performance of the Adaptive Proportion Test over the bit-stream generated with Random Power!'s QRNG is investigated across a spectrum of cut-off thresholds. Notably, results consistently reveal that the measured failure probability remains bounded within the expected 3σ range of the Binomial Cumulative Distribution.

Random Number Generators

Randomness can be considered a core component of security systems, being an essential ingredient of many cryptographic algorithms and protocols. There is clearly a strong link between the capability of ensuring privacy and the ability to generate good random numbers.

At theoretical level, the concept of **random oracle** is used in many proofs and constructions of cryptographic primitives and protocols. It essentially models a perfect source of randomness, that answers queries with values sampled uniformly from a given range. The oracle has memory, in the sense that it will always provide the same answer to the same query. Unfortunately, such idealized objects are not realizable in practice.

Cryptographers turned to realizable ways of providing random bits, the so-called pseudo-random number generators (**PRNGs**). There are many PRNG constructions, some being of mainly historical interest; some are very fast but provide bits with known biases or correlations. Among them, we find algebraic generators, LFSRs, very fast generators such as XORSHIFT based constructions. Mersenne's Twister is one of the most used PRNG constructions. The problem of this class of efficient algorithmic generators, is that it is easy to find relationships between different values produced in sequence. While this can be ignored for certain applications, for security applications this can be a serious problem: even a small bias in the produced bits can lead to full compromise of security [1].

For this reason, cryptographically strong PRNGs (**CSPRNGs**) are typically employed in security applications. Every CSPRNG has a cryptographic algorithm at its core, often a block cipher such as the AES algorithm; this is used to expand an initial seed into a long stream of pseudo random bits, maintaining and evolving a small internal secret state. These objects are essentially randomness amplifiers; they are so good at this, that the produced stream looks perfectly random and would pass any conceivable statistical test. However, there is a catch: anyone knowing the initial seed or the CSPRNG state can predict all following values; the size is equivalent to that of a single key. Thus, several questions arise: who provides the initial seed? What happens if it is compromised? What happens if we want to generate very long sequences of bits?

The answer is that to be safe, the initial seed must be truly randomly generated, and we'll also have to inject fresh entropy in the CSPRNG periodically. This leads to the concept of **TRNGs**, true random number generators that can provide and use fresh unpredictable randomness when needed. Usually, a TRNG includes a randomness source and a PRNG as components, such as in the FIPS constructions standardized in the 800-90 series of documents.

How is true randomness generated? The only possibility is to sample it from the physical world. This may seem obvious; everyone has tossed a coin once to make a tough decision. But in today's connected world we need more secure and more performing ways of doing it. Some TRNGs, for instance in OSs, collect entropy from keystrokes, timing of some external events, disk access patterns. This is a cheap way of doing it because it does not need a dedicated randomness source. However, these methods can easily be tampered. Another possibility is to digitize noise coming from the environment (thermal noise, radio or even audio). It is very hard to demonstrate that such entropy is good, not altered by the sampling method and not alterable by attackers.

Since today's devices are powered by silicon chips, it is natural to build sources of randomness directly in silicon (**on-chip**). A common technique is to use circuits that exhibit and accumulate random timing patterns, for example free-running oscillators.

A new category of TRNGs relies on phenomena governed by Quantum Physics, irreducibly unpredictable by the very same laws of Nature: these objects are called **QRNGs**, arguably the most advanced type of TRNGs. QRNGs are today embodied in a few commercial solutions with different characteristics in terms of exploited properties, complexity, cost and performance. By employing a QRNG, it could be in principle possible to get rid of algorithmic PRNGs, obtaining perfect randomness from its primary source, quantum mechanics.

This type of solution is not attackable by computers, classical or quantum based, giving the user an important tool to reach an absolute level of security. Of course, a QRNG alone is not sufficient to ensure security of a privacy application; we'll talk about other important aspects in the following sections.

Key Takeaways

- Privacy applications need true sources of randomness, algorithmic generators alone are not sufficient to ensure security.
- Quantum mechanics is the primary physical source of randomness.
- A QRNG which can provide perfect entropy and can be easily integrated on silicon is a fundamental pillar for ensuring the highest level of security.

Attacking TRNGs

TRNGs are used to generate cryptographic keys and nonces: we can consider them as the source of all secrets in a security system. Therefore, they naturally constitute a primary **target** for hackers and attackers. The examples are countless; biases, bugs, errors or weaknesses in implementations of TRNGs have been exploited to steal secrets, keys, and money.

As an example, we can cite the case of the **Randstorm** [2] attack; many crypto-currencies products in the time frame 2011-2015 used a weak Java library to generate randomness. Both the entropy collection and the PRNGs were weak, resulting in medium-complexity attacks possible against millions of crypto-wallets. The attack has been unrolled in 2022 and we are steal dealing with its consequences.

Certainly, a **hardware-based** TRNG is more difficult to attack than SW-based algorithmic generators. However, this depends on the actual way in which randomness is sampled and on the physical characteristics of the TRNG.

A very common way of implementing TRNGs on common CMOS processes is to employ free-running **oscillators**. Such circuits, when enabled, oscillate at a frequency determined by the design and the chip fabrication process. Their oscillation is however not fully deterministic: electronic noise affects it, and transistors exhibit random variations of their switching speed, coming from the underlying physics of semiconductor junctions (current is a flow of electrons, and as such exhibit quantum behavior at very small scales). These effects accumulate and macroscopically give an uncertainty in the oscillation phase; if we count the number of oscillations and this uncertainty exceeds the resolution of the counter, we have effectively generated one random bit.

Other more complex constructions have been proposed, such as Transient-Effect-Ring-Oscillators (TEROs), or more complex architectures where oscillators are sampled by other oscillators or combined with linear or non-linear sampling and feedback functions. Besides the difficulty of describing a clear stochastic model for such constructions, they also increase the surface of attacks.

Ring-oscillators based TRNGs are prone to several **attacks**: frequency locking of the oscillators, manipulations of the voltage supply, temperature variations, clock glitches, electromagnetic fault injections, sinusoidal electromagnetic waves injection via power or communication cables attached to the target device embodying the TRNG [3]. Most often, the effect is to lock the oscillators in a predictable behavior, reducing significantly (or even eliminating) the entropy from the produced bits.

All these possibilities of attack come from the fact that, while the basic source of randomness in ring oscillators is coming from quantum effects, the uncertainty is accumulated in a **macroscopic** quantity which is the output of a CMOS circuit, and thus becomes non-quantum in its nature. Its behavior and its sampling can therefore be easily affected by lab equipment. Quantum computers can also probably model these circuits (research in this direction is ongoing) and thus in a post-quantum scenario we cannot guarantee the robustness of these solutions.

Key Takeaways

- A TRNG is the source of all secrets and a natural target of attacks in a security system.
- Classical oscillator-based TRNGs are prone to several categories of attacks.
- On-chip QRNGs are immune to most attacks and tampering techniques that affect classical TRNGs.

Roots of Trust

Good TRNGs are essential components of a security solution, but they are not enough to build up useful applications.

Broadly speaking, we can consider a typical scenario where several electronic devices are manufactured, provisioned with applications and data, and deployed in field; physical access to the devices will then no more be possible, or only at high cost. However, it is mandatory for the owner of the application to keep control of the devices, to remotely monitor their status, verify that they are running the correct version of the software, update it if necessary and that they have not been modified or hacked; through such secure endpoints, monitoring and control can then be extended and exercised onto sensors, actuators, machinery, and other environments. Devices can connect securely to a central server and/or communicate with each other in more complex patterns.

This scenario models well the case of generic IoT applications and depending on the complexity and cost aspects we can range from consumer applications to control of high value assets; the contexts can vary from smart home and building to industrial IoT and to critical infrastructure and power grid monitoring.

What are the security features we need for such devices? An essential one is to have a **trusted identity**. Devices that can be spoofed, cloned or reprogrammed do not bring any value for security applications. We need to uniquely address them, and we need their identity to be strongly bound to the silicon, so that it is not possible for hackers to modify it with logical or physical attacks. Unique identity can be established by provisioning IDs and keys at chip manufacturing or device manufacturing stage or even later on field (such flexibility is more and more required by users).

Leveraging the unique identity of the devices, different rights and capabilities can be assigned; in a security application, the capability of performing a task is almost always linked to the fact of having the correct key. The key is used by cryptographic algorithms and protocols to authenticate and decrypt messages and data. Keys can be generated on the device or assigned externally by the owner of the application; in any case, they must be generated with good TRNGs. Keys must also be managed and stored in a secure way, so that it is not possible to extract them from the device and/or transport them to other devices; thus, **key generation**, **key management and key storage** are essential security features.

Secure communications are another fundamental pillar of a security application. We must be able to reliably communicate with devices, authenticate them and be sure that no one can spy or alter the communication flow. Many different types of communication protocols exist, depending on the complexity and requirements that we set on our network. However, all of them require fresh and reliable randomness, for instance to avoid messages to be replayed by attackers or to compromise secrets in the long term. More and more often, the requirement that the content of the communication should not be accessible by quantum computers is set; this demands the usage of so-called **post-quantum cryptography (PQC)**.

These security features we outlined can be grouped under the definition of a **Root-of-Trust** (RoT) [4]. A RoT, correctly designed and implemented on a silicon chip, enables all kind of security applications, and enforces the rights and control of the application owner on the device. Roots of trust can be designed and implemented in several ways, as stand-alone components or as parts of more complex embodiments such as system-on-chip (SOCs). We can generally distinguish different types of RoTs:

- **Discrete component:** in this case the RoT resides on a dedicated **ASIC**, that is typically tamper resistant and certified according to well-known security standards (e.g. FIPS 140-3, or common criteria). This discrete component can be a TPM chip, a secure element, or a more complex microcontroller-based IC; it contains persistent memory to store critical assets such as master keys and storage keys (keys can be generated on-chip or imported from other hardware security modules). Most of the RoT is fixed in hardware but there may be the possibility to update part of the software running on the discrete element. This solution features a **very high level of security** and availability but adds an overhead for the device manufacturer; the cost of adding a chip to a device can be justified if the features provided are critical for the application.

- **Integrated hardware IP:** in this case the RoT is still implemented as dedicated hardware, but it is integrated as part of a larger chip, typically a SoC. The SoC can run complex applications, manage different communication stacks and user interfaces, while the RoT is the main provider of security. The RoT storage on chip is typically implemented as a reserved part of the chip memory, shared in part with the application CPU. This solution brings the advantage of a **high security level** and that of a **reduced price**; silicon surface is still accounted for the RoT, but there is no need for an additional chip. Examples are integrated SIMs and integrated secure elements. Integrated RoTs may suffer from some limitations due to the limited amount or re-programmability of on-chip memory, and therefore are typically less complex that discrete implementations.

- **Firmware:** in this case the RoT is implemented with code that is loaded during the start-up of the system and is executed with a certain degree of isolation from the main application. Examples are RoTs that are executed in a trusted-execution-environments (as secure software components or trusted applications), or even in re-programmable logic in case the SoC is an FPGA (in this case isolation is typically stronger). We can say security is at generally at good level, depending on the SoC features; this type of RoT does not create significant additional costs for the device, which only has to allocate the correct amount of resources to the RoT (memory, code size, execution time). Firmware RoTs are also **easier to update** and fix compared to hardware ones.

The correct type of RoT for a given device and application can be selected basing on considerations such as the added price, the level of security and the flexibility; there is no best solution in general. If the RoT employs method and algorithms that are resistant to the advent of quantum computers, we have an instance of a Quantum-safe RoT (**QRoT**). A QRoT that makes use of an on-chip QRNG is probably the most advanced type of RoT we can imagine, giving us ultimate security level.

Key Takeaways

- A good TRNG, while essential, is not sufficient to build a full security application.
- A root-of-trust (RoT) enables a device with all needed security primitives: TRNG, trusted identity, key generation and management, secure communications.
- RoTs can be implemented in different ways, as ASICs, integrated in silicon or as firmware components to best fit the customer requirements.
- A quantum-safe RoT (QRoT) is resistant to quantum computers and can employ perfect randomness coming from a QTRNG to build up the ultimate hardware security solution.

Privacy Applications

Good randomness is important in the context of connected devices and infrastructures, and in general to enforce data confidentiality and authentication in machine-to-machine communications but is also becoming crucial to guarantee privacy of human users in the digital world.

As more and more services are run over the internet, our data is potentially stored and used in a multitude of ways by different private companies, governments, institutions: most concern is on medical data, tracking of our position, other personal data like shopping preferences and habits, financial transactions, etc.

Privacy enhancing and privacy preserving technologies (**PET**) are a class of techniques that help, at different degrees, reaching the level of privacy required by national laws and by data regulations. This is a broad class of protocols and algorithms, among which we can cite the following examples:

- Secure multi-party computations (**MPC**), where a secret is split in different parts among several users; they can then collaborate to compute a function over their collective private data without anyone ever revealing his share to others. There is an increased interest in related financial applications (for example crypto currencies and **CBDCs**).
- Zero-knowledge proofs (**ZKP**) can be used by users to prove knowledge of a piece of information or of a correct execution of a program without revealing anything about it.
- Oblivious technologies are helpful to hide the way in which data is accessed and allow users and data analysists to access remote databases without revealing the utilization patterns and content. Oblivious Random Access Memories (**ORAMs**) are an interesting primitive that is recently seeing a sparkle of interest and evolution.
- Fully Homomorphic Encryption (**FHE**) is used to allow analysis to perform arbitrary computations on data which remains encrypted for the whole duration of the process.
- Differential privacy can help with the anonymization of data and can guarantee that identity of single users cannot be revealed when their data is aggregated into larger data sets.

All these techniques make use of **advanced cryptographic algorithms** proposed and investigated by academia and private companies, with the aim of rapidly advancing the state of the art and allowing privacy to be preserved in a growing set of complex scenarios. We can comprise such cryptographic techniques under the broad definition of advanced cryptography [5], to distinguish them from simple cryptographic primitives.

Advanced cryptography is more complex, much slower to execute and harder to deploy than basic cryptography but can address use cases that were not even conceivable only few years ago; the gap is also shrinking due to intensive research and development of the field and a growing effort of standardization. Advanced cryptography needs good randomness sources as a core service and can miss the goal of privacy preservation if fresh randomness is not available at sufficient bitrate.

Key Takeaways

- Users' privacy is rising as a primary concern in the digital world, and this impacts many sectors such as healthcare, financial and institutional.
- Privacy enhancing technologies built on advanced cryptography can help ensure privacy.
- Advanced cryptography is complex and under continuous development: it needs a good randomness source at its core.

Standardization and Compliance

Today's world is more interconnected than ever, and globalization is removing many boundaries against the diffusion of technology and the spreading of innovative solutions across all countries; companies strive to keep the pace and reach customers everywhere in the world to sieve growing commercial opportunities. Internet-of -things is becoming a reality, with the goal of making our life easier and richer in experiences.

An essential aspect of this quickly evolving scenario is the interoperability of systems and devices; while fragmentation can be expected during initial phases of commercialization, deep market penetration must rely on established standards. Products that comply with international norms have a higher chance of gaining widespread diffusion.

Standards can relate to different aspects of a product, for instance its functionality, its safety, its security, its interoperability. They can be emanated by **international standard bodies** (e.g. ISO) and/or by national entities. Due to the political and commercial relevance, US and EU standards have a well-established reputation, and have de-facto become worldwide standards.

As we have already discussed in the previous sections, secure electronic systems rely on cryptographic algorithms and protocols at their core to provide users with privacy guarantees. It is thus understandable how the usage of standard cryptographic techniques is not only suggested, but mandatory to ensure robustness and ensure trust in a security solution.

NIST, the US institute for standards and technology, maintains a large library of publications pertaining all aspects of cryptographic algorithms, cybersecurity practices and implementations. For example, the publications series 800-90 standardizes the design and validation of TRNGs [6]. NIST runs a program for validation of cryptographic modules called **CMVP**; products certified under this framework gain the label of FIPS 140-3 certification, which brings additional value to a commercial device and is even mandatory for selling the product in some market segments, such as defense and homeland security.

Several other security certifications schemes exist and are referenced worldwide, notably the **Common Criteria framework**, the **SESIP** certification by **Global Platform**, the **PSA** certification for IoT products initiated by **Arm** [7], the **EMVCo** certification for payment systems and others.

The standardization of advanced cryptography is today an open problem, due to the complexity of the algorithms involved; it is expected that related specifications and norms will be established soon. It is essential that security products stay up to date regarding such processes.

NIST National Institute of Standards and Technology U.S. Department of Commerce

FIPS 140-3™

Key Takeaways

- Standardization and interoperability of products is not only desirable but is now a must in the globalized economy.
- US certification path established by NIST is a worldwide de-facto standard for cryptographic devices.

QRNG and supply chain

The global economy is also highly affecting the manufacturing of integrated circuits, which are at the core of many devices that have become so important in our daily life. It is a fact that most of the **ASICs** are produced in the far east region, thus not only touching the economic sphere but also the geopolitical equilibrium between nations. The level of investments needed to be at the cutting edge of chip manufacturing is today measured in tens of billions of dollars, allowing only few players on the planet to enter the game; this reflects also on the design side: it is estimated that the tape-out of the M3 chip has costed Apple around one billion dollars overall.

On the other hand, chips are necessary to build today's complex communication infrastructures, our smartphones, cars and almost every electronic object with which we interact. It is thus understandable that governments are trying to exercise control over their design, their availability and their usage, asking chip designers to comply with the local laws, which sometimes require difficult trade-offs with users' privacy.

The scenario is obviously even more involved when part of the supply chain is located in a different country. At academic research level, it has been shown how in principle stealthy **hardware trojans** [8] can be introduced in silicon during the manufacturing process, thereby creating a gap in the chain of trust. For these and other reasons, states are putting more effort in defining strategies for keeping sovereign control on the integrated circuits supply chains, by promoting acts such as the US or European **Chips Act**.

Control over the supply chain does not only mean enforcing the integrity of the manufacturing stage, but also ensuring that the ASICs are available when needed, in sufficient quantities and that can be produced for a sufficiently long time as to permit maintenance and replacement of already deployed systems, but also the business continuity of device integrators.

This is even more true for ASICs that provide advanced functionalities, such as those that can accelerate AI applications or machine learning and make it more ubiquitous. Speaking about cryptographic technologies, we see an increased interest on post-quantum cryptography (**PQC**) [9], as the race for quantum computing is also a battle front for several nations.

As QRNGs are on the forefront of randomness generation, they also constitute an important asset for those systems that need to integrate the highest possible level of security. A **QRNG** technology which is easily manufacturable and that can provide certainty about its supply chain has clear advantages over competing solutions.

Key Takeaways

- ASICs supply chain has become a key factor in economic and geopolitical terms.
- Advanced solutions that are on the forefront of technology but are also easy to manufacture and integrate give a clear strategic advantage.

References

1. Joachim Breitner, Nadia Heninger, Biased Nonce Sense: Lattice Attacks against Weak ECDSA Signatures in Cryptocurrencies, Proceedings of FC 2019, pp 3–20, available at →https://eprint.iacr.org/2019/023.pdf
2. →https://www.unciphered.com/blog/randstorm-you-cant-patch-a-house-of-cards
3. [3] A. Theodore Markettos, Simon W. Moore: The Frequency Injection Attack on Ring-Oscillator-Based True Random Number Generators, Proceedings of CHES 2009, pp 317–331, available at →https://www.iacr.org/archive/ches2009/57470316/57470316.pdf
4. →https://trustedcomputinggroup.org/
5. Shai Halevi: Advanced Cryptography: Promise and Challenges, Proceedings of ACM CCS '18, →https://www.youtube.com/watch?v=tsQ5Bf8KA6E
6. NIST: Random Bit Generation, →https://csrc.nist.gov/projects/random-bit-generation
7. →https://www.psacertified.org/
8. Shivam Bhasin, Francesco Regazzoni: A survey on hardware trojan detection techniques, Proceedings of IEEE ISCAS 2015
9. →https://csrc.nist.gov/projects/post-quantum-cryptography

The Random Power consortium for the EU Attract Project:


Contacts:

Random Power s.r.l.

Registered Office:
www.randompower.eu
Via Macedonio Melloni 40,
20129 Milano - ITALY