PKWARE

pkware.com

The Entropy Problem

Random Data and Secure Cryptography

2021

Contents

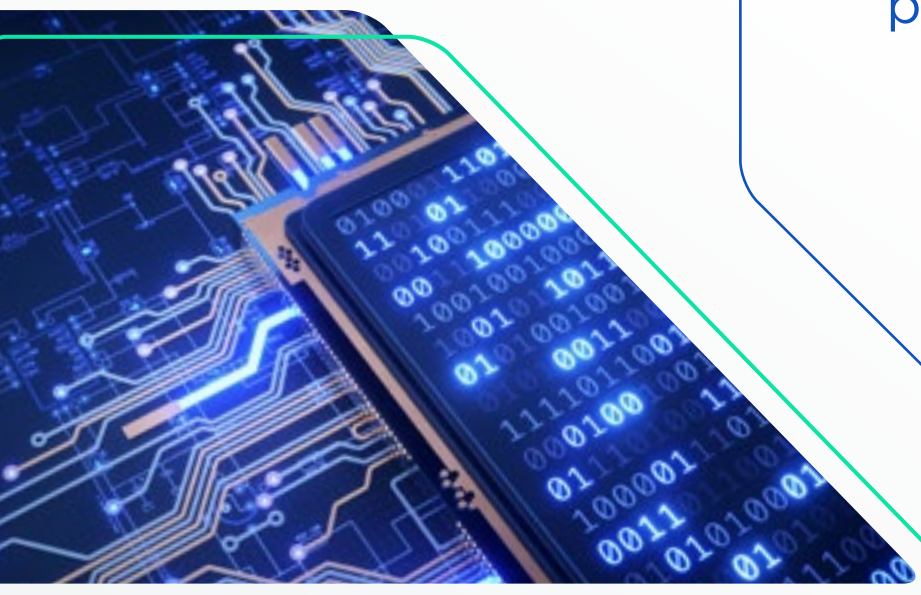
Executive Summary

The strength of any cryptosystem depends in large part on the unpredictability of the data used in the encryption process. Unfortunately, some of today's most commonly used sources of "random" data depend on inputs that have the potential to inject predictable data, and therefore weakness, into the process.

Low-entropy data sources produce encryption keys that can be attacked much more easily than a truly random key. Even high-performance pseudorandom number generators (PRNGs) that have been certified as "cryptographically secure" may prove to be insufficiently random once large-scale quantum computers become available. Full-entropy random data provides the highest possible security against potential key attacks. Even quantum computers, while they may be able to break the asymmetric keys currently used in public key infrastructure, are expected to be ineffective against truly random AES-256 encryption keys.

Random number generators (RNGs) that measure quantum physical processes are able to deliver truly random data at speeds up to 1 Gb/second, effectively solving the entropy problem for government entities and other organizations that store and process highly sensitive information.

Full-entropy random data provides the highest possible security against potential key attacks.

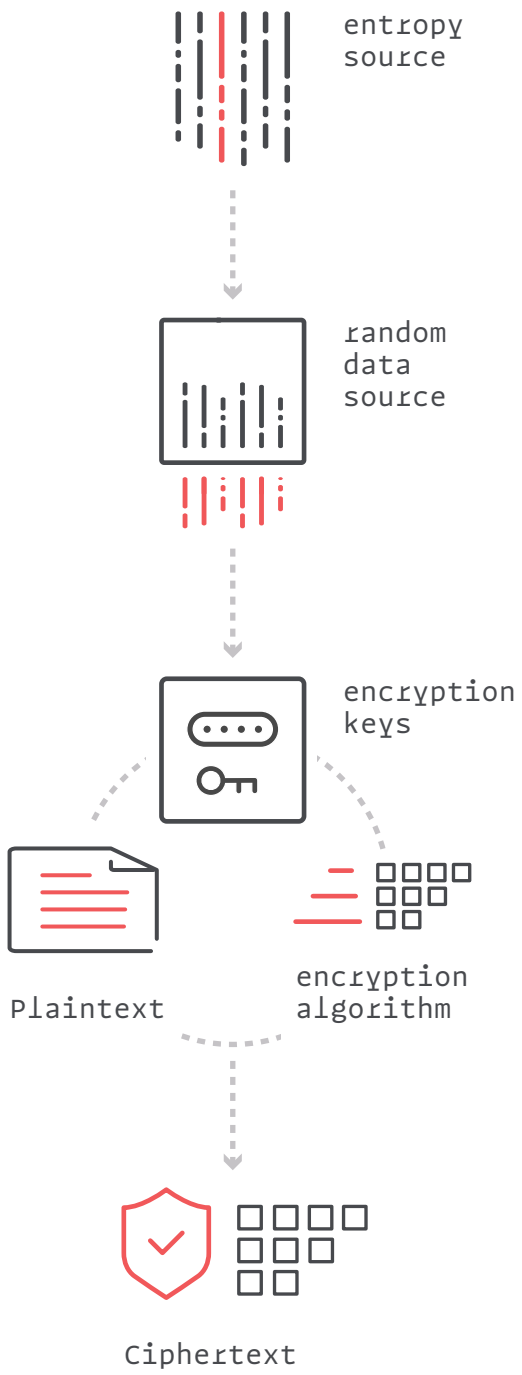


Why Random?

Together with strong encryption algorithms and secure key management practices, random data is a foundational element of cryptography. Random numbers are used as inputs for key generation, key wrapping, authentication, and many other common cryptographic functions.

The randomness of data is described in terms of its entropy, or unpredictability, which can be measured using statistical tests. Higher entropy means that data is less predictable and therefore more suitable for use in cryptography. Data with an entropy value of zero is completely deterministic, while data with an entropy value of one (full entropy) is considered to be truly random. Full-entropy data makes it impossible for anyone to predict a given bit in a data stream with more than 50 percent accuracy.

Without full-entropy random data, cryptography is unable to deliver its promised level of security. The AES-256 encryption algorithm, for example, is considered strong enough to render brute-force attacks useless, as even the most powerful computers would require billions of years to try all possible keys. However, the inclusion of predictable data in the key generation process can create shortcuts that make attacks much more likely to succeed.



Unpredictable data is the raw material from which encryption keys are made. Any flaw in the initial source of entropy will be carried through to the end, resulting in potentially vulnerable ciphertext.

Weaknesses in Common Approaches

Random data for cryptographic applications is typically obtained from a physical RNG, a software-based PRNG, or from a combination of the two.

These technologies, when properly implemented, are able to pass standard tests for randomness and cryptographic security. However, most of today's common approaches to random number generation have limitations that can leave sensitive data vulnerable to attack.

Physical RNG

Hardware-based RNGs produce random data by "collecting entropy," meaning that they measure events that are expected to be random. Entropy can be collected from the external environment (using phenomena such as ambient sounds or even cosmic background radiation), or from within a computer (using events such as hard drive activity, voltage fluctuations, or keyboard and mouse interactions).

One area of concern regarding physical RNGs is the possibility that the events being measured could be manipulated in order to produce predictable output. This scenario, while unlikely in most environments, cannot be ruled out entirely when extremely sensitive data is in question.

A more practical concern is that most physical RNGs produce data at unacceptably low throughput (number of bits generated per second) due to limitations in the phenomena being measured. Organizations relying on these devices must compromise on either the level of entropy in their data or the speed at which cryptographic functions can be completed, either one of which can put sensitive information at risk.

Today's common approaches to random number generation have limitations that can leave sensitive data vulnerable to attack.

PRNGs

PRNGs (also known as deterministic random bit generators) are software-based systems used to produce higher-throughput data at entropy levels that are good enough for many common cryptographic uses.

A PRNG uses an algorithm into which an initial seed value is fed, in order to define the generator's state. The algorithm then performs a series of operations using the seed value and generates a stream of data much longer than the seed itself. Depending on the implementation and purpose of the PRNG, the seed value might come from a physical RNG, a table of predetermined values, or another source.

A large amount of research has gone into creating reliable PRNGs. Some approaches have been certified by the US National Institute of Standards and Technology (NIST) as being "cryptographically secure" and acceptable for use in high-security settings. Nevertheless, it is important to note the limitations of PRNGs in general:

- Deterministic design: A PRNG has no intrinsic entropy and can never produce truly random data. The algorithms used are deterministic by nature, so a given seed value will always produce the same output.

- Potential for hidden defects: PRNGs using outdated or poorly designed algorithms generate predictable data. However, flawed algorithms are often difficult to identify until it is too late, when the weakness has been exploited.

- Implementation issues: Even cryptographically secure PRNGs are dependent on proper configuration and implementation in order to function properly. As with flawed algorithms, improper implementations are often difficult to identify until after a vulnerability has been used in an attack.

- Vulnerability to compromise: The algorithms used in PRNGs are susceptible to intentional weakening. The Dual_EC_DRBG algorithm, for example, was in widespread use until 2014, when it was removed from NIST guidance because of a backdoor reportedly inserted by the NSA.

Running Low on Entropy

PRNGs that depend on system information as an entropy source can encounter performance issues during and shortly after the device has started up, when system activity is relatively predictable and user activity is lower than normal.

PRNGs running on virtual machines face an even greater challenge as they often lack direct access to information on the system activity or user interactions that could be used to populate their entropy pool. Furthermore, if PRNGs are running on multiple machine images that were created with the same initial state, they are likely to produce identical output.

Potential Attacks

Due to the limitations of PRNGs and traditional physical RNGs, information encrypted using these technologies may be vulnerable to a variety of attacks. Attacks can take many forms, but they typically follow one of the strategies listed below.

- Analysis of PRNG output: Attackers can evaluate the data stream produced by a PRNG and look for patterns, which can be analyzed and used to decrypt protected information. This is generally unfeasible against a cryptographically secure PRNG, but is a significant concern with weaker algorithms or flawed implementations.

- Knowledge of PRNG inputs: Because all PRNGs are deterministic, knowledge of the seed will allow an attacker to reproduce the generator's output. When a PRNG is seeded from a low-entropy source, hackers may be able to guess the seed value with relative ease. This vulnerability gained attention in the early days of the internet, when a low-entropy PRNG allowed hackers to decrypt Netscape's SSL-encrypted traffic using only consumer-grade technology.

- Manipulation of PRNG inputs: If hackers are unable to guess a PRNG's seed value, they may be able to control it instead. Tampering with the inputs of a physical RNG may be challenging or impossible, but cybersecurity experts have identified several methods of manipulating the inputs of a PRNG.

In the event that a PRNG algorithm is intentionally weakened, other forms of attack become unnecessary. Anyone with a knowledge of the backdoor would be able to use it to decrypt and exploit information that was encrypted using the compromised algorithm.

The risk of a successful attack can be decreased through the use of cryptographically secure PRNGs, but it cannot be eliminated completely. As hackers gain access to more sophisticated tools, even the most secure PRNGs may prove to be ineffective.
The risk of a successful attack can be decreased through the use of cryptographically secure PRNGs, but it cannot be eliminated completely.

Quantum Random Number Generation

Until recently, cryptographically secure PRNGs were the only practical source of unpredictable data for use in encryption. While some physical RNG technology was able to provide full entropy, it could not deliver the throughput needed for enterprise-scale applications. Organizations were forced to accept the trade-off of lower entropy in order to gain the volume of random data they required.

Today, however, developments in the field of quantum random number generation have made this compromise unnecessary. Quantum random number generators (QRNGs) can produce full-entropy random data at speeds of up to 1 Gb/second, equivalent to the output of the highest-capacity PRNGs and enough to meet the needs of even the largest organization.

QRNGs detect random quantum effects and convert those fluctuations into a stream of binary digits. As quantum phenomena are random by definition, the data generated by a QRNG has full entropy and cannot be predicted by any means.

Output from a QRNG can be used for key generation or any other cryptographic use, without the need for an external seed or other potential vulnerability. This approach eliminates the trade-offs associated with other random data sources and provides the highest possible security against potential key attacks.

Quantum Computing vs. Quantum Number Generation

Quantum random number generation is an established technology that measures unpredictable quantum phenomena to generate truly random data.

Quantum computing is a rapidly developing, but still largely theoretical, approach to information processing within a computer. Rather than using binary bits, which can store only two possible values (zero or one), quantum computers will store information using quantum bits (also known as qubits), which have the potential to encode more information per bit and allow for faster and more complex processing.

Quantum computers, when they are available for large-scale use, will likely be able to perform difficult tasks such as integer factorization quickly enough to break most of the asymmetric encryption in use today. However, AES-256 and other forms of symmetric encryption are expected to remain secure against quantum computers.

Other new technologies based on quantum physics, such as quantum key distribution, may also provide security against the threat of attacks from quantum computers.

PK Protect

Enterprise Key Management With Quantum Random Number Generation

PK Protect Enterprise Manager appliances combine enterprise-wide encryption key management with full-entropy quantum random number generation, providing a unified solution for protecting highly sensitive data.

PK Protect appliances are available in four versions, including a virtual appliance and three hardware configurations.

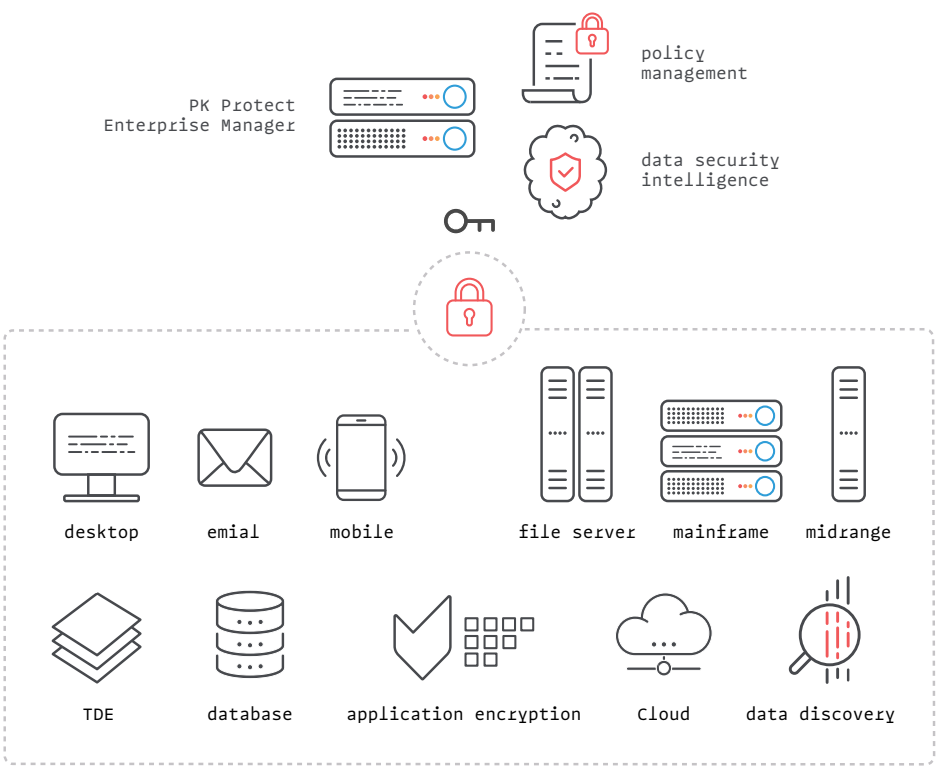
- PK Protect Enterprise Manager 200v, a virtual appliance
- PK Protect Enterprise Manager 300h, with a FIPS 140-2 Level 3 hardware security module (HSM)
- PK Protect Enterprise Manager 300h, with a FIPS 140-2 Level 3 hardware security module (HSM)
- PK Protect Enterprise Manager 350, with both the HSM and RNG

Enterprise-Wide Encryption Management

The PK Protect Enterprise Manager is the central component of the PK Protect platform, allowing administrators to define an organization's data protection policies and apply them across the entire company.

When configured with an optional QRNG, the Enterprise Manager delivers true random data for use in key generation and other applications. Further details on the PK Protect QRNG are provided on the following page.
In addition to full-entropy random number generation, the PK Protect Enterprise Manager offers several other unique features:

- Support for persistent encryption, which remains with data wherever it is shared or stored
- Out-of-the box support for encryption on endpoint devices such as desktops, laptops, and phones (Windows, Mac, iOS, and Android)
- Support for automated data discovery and remediation
- Automated key management that eliminates the complexity of key generation, exchange, synchronization, and rotation

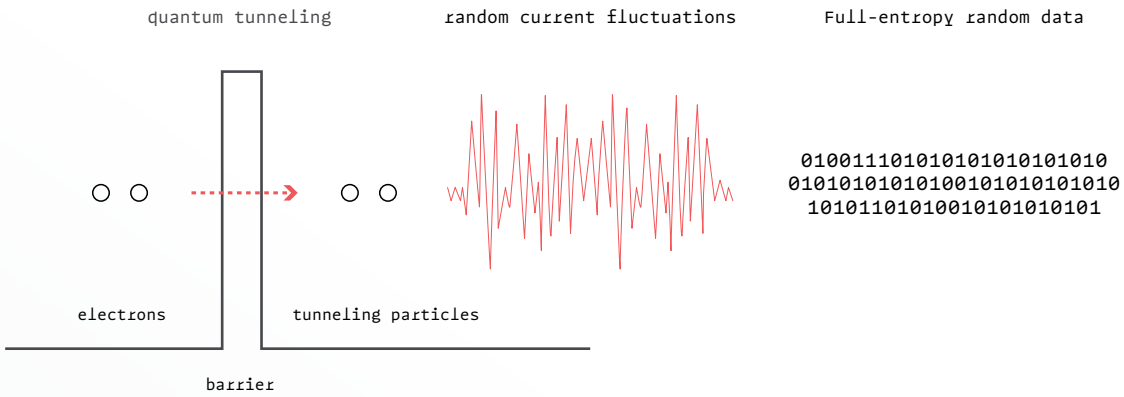


Full-Entropy Data at Unprecedented Rates

The PK Protect QRNG was developed by QuintessenceLabs, PKWARE's technology partner and the leading provider of quantum cybersecurity solutions.

The QRNG generates random data by measuring quantum tunneling noise. Quantum tunneling is a phenomenon in which a particle travels across a barrier that—according to classical mechanics—it should not be able to cross.

Inside the QRNG, a voltage is applied to a forward-biased diode junction. The diode contains a barrier through which charge carriers can "tunnel," even if they lack the energy to overcome the barrier according to Newtonian physics. The number of particles that will tunnel through in a given instant in time cannot be predicted, making the process an ideal source for random data.



Quantum tunneling in the diode creates random fluctuations in the current flowing through the diode. These fluctuations are measured, digitized, and digitally processed to generate ultra-high bandwidth random numbers. Full-entropy data is generated at 1 Gb/second, suitable for use in any cryptographic application.

About PKWARE

PKWARE offers the only data discovery and protection solution that locates and secures sensitive data to minimize organizational risks and costs, regardless of device or environment. Our ultra-efficient, scalable software is simple to use on a broad range of data types and repositories, enabling precise, automated visibility and control of personal data, even in the fastest-moving, most complex IT environments. With more than 1,200 customers, including many of the world's largest financial institutions, retailers, healthcare organizations, and government agencies, PKWARE continues to innovate as an award-winning global leader in data discovery, security, and compliance. To learn more, visit PKWARE.com.

Enterprise-Wide Policy Management

The PKWARE Enterprise Manager provides a single point of control for data protection activity across the entire organization

Simple Workflow

With PKWARE, data protection is automated for end users and easy for administrators to manage

Easy Implementation

PKWARE supports a variety of deployment options, enabling organizations to implement their data protection solution without time-consuming changes to infrastructure and workflows

Protection Without Gaps

PKWARE works on every enterprise operations system and provides persistent protection that remains with data even if it's copied or shared outside organizations

Integrated Discovery, Classification, and Protection

No other solution has the capability to find, classify, and protect data in a single automated workflow

Multiple Protection and Remediation Options

Organizations can take a policy-based approach to data protection and choose from action including persistent encryption, quarantine, masking, and deletion.

BOOK A DEMO

PKWARE.com

866-583-1795

201 E. Pittsburgh Ave.
Suite 400
Milwaukee, WI 53204