White Paper

Enabling GDPR Compliance Through Innovative Encryption AND Key Management Approaches

Sponsored by: Hewlett Packard Enterprise

Duncan Brown
August 2016

IDC OPINION

GDPR is the biggest shake-up in European data protection legislation for 30 years. Organizations have less than two years to ensure that their data protection processes are compliant. Most organizations will struggle to meet the May 25, 2018 deadline.

Although GDPR is not prescriptive in the technologies required to enable compliance, it strongly hints at the use of encryption and pseudonymization as approaches to protect sensitive data. There are three main reasons why these technologies receive particular attention in the text of GDPR:

- Encryption can be used to mitigate the risks inherent in data processing, such as unauthorized disclosure of, or access to, personal data.
- The requirement to notify data subjects (such as consumers or employees) of a data breach is removed if the data is rendered unintelligible using a measure such as encryption.
- The use of pseudonymization can reduce the risks to data subjects while helping data controllers and processors meet their compliance obligations by minimizing both the exposure of personal data and the opportunities to identify data subjects.
- Although not explicitly stated in GDPR, the extensive rules around data transfers outside the EU can be simplified by encrypting the data and managing keys in-country (including on-premise).

Many organizations are therefore examining encryption and pseudonymization technologies; however, they quickly discover the complexities and management overheads with traditional approaches.

HPE Format-Preserving Encryption (FPE) allows a highly granular approach to encryption, which facilitates field-level protection. It also preserves business functionality, meaning that normal data processing activities are maintained even though the data is encrypted. FPE fulfills both encryption and pseudonymization functions, which makes it a particularly useful technology in the context of compliance with GDPR.

One of the complexities introduced by encryption is the management of keys. Few organizations have the skills and processes required to manage encryption keys, particularly at high volumes. HPE Stateless Key Management is an approach that minimizes this overhead by generating keys on demand, rather than having them stored centrally.

Encryption should not be a "one size fits all" activity, and other approaches should be regarded as complementary. For example, email encryption, such as HPE SecureMail, protects personal data and ensures that it can only be read by the intended recipient. Infrastructure management solutions, such as HPE Enterprise Secure Key Manager (ESKM), provide a centralized key management hardware-based appliance solution for unifying a key repository and automating an organization's data-at-rest encryption policy.

There are many aspects to complying with GDPR, but IDC believes that innovative approaches to encryption and key management, such as HPE FPE and HPE Stateless Key Management, substantially enhance and simplify any organization's compliance activities.

DISCLAIMER: IDC does not provide legal advice. If in doubt, consult a lawyer

WHY ENCRYPTION IS IMPORTANT TO GDPR

Introduction to GDPR

GDPR is a welcome and long overdue refresh of Europe's data protection laws. It replaces current legislation that dates to 1995, pre-dating the dotcom boom, Twitter, Facebook, and the cloud. GDPR updates the law to account for these and future developments that create and use personal data. An additional benefit of GDPR is that it applies to all EU member states, rather than the hotchpotch of differing regimes that exists in the region today.

After a tortuous four years of discussion and debate, GDPR was signed into law in April 2016 and becomes effective on May 25, 2018. We have, therefore, less than two years to ensure compliance. And comply we must: the penalties for non-compliance could reach 4% of global annual revenue or €20 million, whichever is the greater. GDPR also introduces mandatory breach notification, the consequences of which concern executive boards that worry about reputational damage.

Note that the UK's prospective exit from the EU does not materially affect the broad GDPR landscape:

- UK firms that process EU citizen personal data will have to comply with GDPR anyway, because GDPR applies extra-territorially.
- The UK is likely to implement local laws similar to GDPR in order to facilitate data transfers from the EU, as governed by 'adequacy' rules.

Data Security Concerns

It is important to note that GDPR is much more than a data security issue. It contains a raft of new measures including data portability, consent and revocation, age verification, and the right to be forgotten. However, a large part of achieving compliance comes down to good data security. What does GDPR say about this specifically?

In fact, GDPR is remarkably non-prescriptive when it comes to defining security measures. Of its 99 articles only one (Article 32) talks specifically about security and it is particularly vague. The regulation mentions the use of pseudonymization and encryption, as well as good security practice, back up, and testing.

There are, however, two key areas of GDPR scope that relate directly to data security. The first of these, and the one that receives most attention in the media, is the transfer of data to so-called third countries (that is, non-EU countries). The EU has taken a very defensive and a specific view on data

transfers, the essence of which is to prevent organizations from exporting data beyond the EU in order to avoid data protection regulation. GDPR, therefore, has an extra-territoriality clause that extends the jurisdiction of GDPR to the processing of EU citizen data "regardless of whether the processing takes place in the Union or not" (Article 3).

This means that organizations exporting data, for example by using cloud services, are subject to the terms of GDPR and must adhere with very specific requirements. In contrast to the solitary article on security, data transfers are afforded an entire chapter (Articles 44 to 50) spanning six pages.

Companies worry about data transfers for three reasons:

- They worry that once data is stored in another country it will not be subject to the governance laws and data protection practices that may then impact the security of that data, as well as compliance obligations set by the EU.
- They flinch at the legal regimes they need to consider, such as binding corporate rules and modelling contract clauses, that their customers may demand.
- They are concerned that data transferred to third countries may be subjected to unauthorized access, either by malevolent actors or by nation states.

For these reasons, many companies in the EU are reluctant to put sensitive data in the cloud. They are also cautious about encrypting data without a strong understanding of key management principles and practices.

Data Residency and Sovereignty Issues

Data residency and sovereignty are concepts that are not specifically discussed in GDPR. The terms relate to data transfers, and in particular the legal jurisdiction that applies to the data. There is a concern, sometimes but not always with a legal basis, that transferred data somehow contravenes existing data protection law. GDPR clears this up and removes differing legal bases and variations that occur across the EU. GDPR is quite clear: data transfers are allowed, subject to the following provisions:

- Adequacy: The EU may decide that the third country has an "adequate" level of data protection, essentially equivalent to that specified in GDPR.
- Binding corporate rules (BCRs) or model contract clauses (MCCs): legal structures that obligate data processors (such as cloud service providers) to abide by the rules specified in GDPR.
- Consent from the data subjects.
- Mutual legal assistance treaties (typically in the cases where a law enforcement agency in a third country requests data disclosure in order to pursue a criminal investigation).
- A number of other processes involving structures yet to be implemented (such as codes of conduct and certification mechanisms).

The judgement of the EU as to the adequacy of third-country data protection laws is an important one. Currently, the EU maintains a list of 12 countries that are deemed to have adequate protection. the countries currently assessed as having equivalent data protection laws are: Andorra, Argentina, Canada (commercial organizations), Faeroe Islands, Guernsey, Israel, Isle of Man, Jersey, New Zealand, Switzerland, and Uruguay.

Importantly, the US is not deemed to have adequate data protection laws, primarily because there is no federal data protection legislation. Instead, the US and the EU have a separate agreement (not legislation) called EU-US Privacy Shield. Privacy Shield was agreed in July 2016 in response to the striking down of its predecessor, Safe Harbor, which was deemed by the European Court of Justice to be inadequate (as revealed by the Snowden revelations). It remains to be seen whether Privacy Shield will survive a similar investigation by the ECJ, as seems likely to take place. In the meantime, interest grows in BCRs and MCCs as providers take unilateral action in order to protect their customers' data, and their business in the EU.

These issues are important to consider because most EU firms want to use US-based providers. Why? Because US providers have the scale that underpins the cost, flexibility, and manageability desired by firms: there are no EU-based providers that can match the infrastructure of AWS, Microsoft, IBM, or Google.

There is another option that removes the need for data transfers at all, and that is to keep the data resident in the EU. Many EU-based firms prefer this model because it removes consideration of transfer rules and whether Privacy Shield or BCRs are required. All the major datacenter and cloud providers have EU-based facilities. However, firms also worry that that data held within the EU can be transferred to another country without their knowledge: this may be due to espionage or because that country forces the datacenter or cloud provider to hand over the data. Such jurisdiction clashes are increasing, notably between the US and the EU. What can firms do to protect their data in this case?

One option pioneered by Microsoft is to use a "data trustee" model, in which an EU-based partner (T-Systems is the first such collaborator) holds physical and logical access to a Microsoft datacenter in Germany; on receipt of a disclosure warrant Microsoft would not be able to comply itself but would have to refer the law enforcement agency to T-Systems via German law.

Another strategy to avoid unauthorized data disclosure is to encrypt the data and hold the keys securely within the EU. This is likely to be a less expensive, more legally straightforward solution. Any data transferred is encrypted and effectively useless to the recipient.

What GDPR Says About Encryption and Pseudonymization

Encryption is mentioned four times in the text of GDPR (Recital 83, Articles 6, 32, 34), and in each case it is positioned as an approach to mitigate risks in the processing of data. Importantly, GDPR positions encryption as a mechanism that renders personal data unintelligible to unauthorized individuals, which is a mitigating action against both a data breach and the requirement to make a public notification of that breach. In other words, encrypted data is not regarded as personal data for the purposes of breach notification.

Pseudonymization is an approach that appears substantially more often than encryption within the text of GDPR (Recitals 26, 28, 29, 75, 85, 156, Articles 4, 6, 25, 32, 40, 89). Pseudonymization is the process of removing personal identifiers from data for the purposes of processing that data while protecting the identity of individuals. The implication here is that companies should be able to process data for a variety of purposes without necessarily being able to identify the individual whose data it belongs to. Companies use personal data for a variety of purposes in which identification is necessary, such as reporting, analytics, and backup. Even core business processes such as order taking can be undertaken without the identity of the customer being revealed.

There is an important distinction here between encryption and pseudonymization: most encryption approaches involve the scrambling of data such that it is unintelligible to unauthorized individuals — and to business applications. Encryption, therefore, often results in a loss of application functionality, which therefore impedes business processes. This is one of the primary reasons why encryption is used very selectively (and often not at all) within organizations.

It is important to note that neither encryption nor pseudonymization are mandated by GDPR: given the emphasis placed on these approaches (particularly pseudonymization) it seems that the EU is giving companies a very heavy hint as to what they regard as best practice. However, it is also important to understand that encryption and pseudonymization are very generic categories of approach, and there are many variants with differing levels of capability, flexibility, and manageability.

Why Use Encryption?

The new EU data protection law is complex. There is a broad range of new requirements with which organizations need to comply, with very little direction or prescription as to how those requirements can be met. Enterprises need to make up their own minds as to how best to reach and maintain compliance.

Of all the possible approaches to GDPR-compliant data security, encryption and pseudonymization have the strongest tangible link to the text of GDPR. There are two reasons why IDC thinks that encryption makes sense for enterprises to utilize in the context of GDPR:

1. Encryption facilitates risk mitigation: encrypted data is deemed to be inherently less exposed to risk (Recital 83, Article 6, Article 32).
2. Encryption renders personal data unintelligible, meaning that notification to data subjects is not required in the event of a data breach (Article 34, section 3 (a))

There are other, business-related reasons to use encryption. These fall into two broad categories:

- Encrypted data avoids accidental loss. Devices such as laptops and mobile phones are easily misplaced, but encrypt the data on these devices and they would not be exposed. In addition, encryption avoids data loss from misdirected emails, shared cloud-stored files and other data loss scenarios.
- Encryption protects commercially sensitive data, such as intellectual property. While GDPR may mandate the protection of personal data there are strong commercial reasons for doing this regardless of the prevailing data protection regime: losing customer data is a fast road to severe reputation damage.

The persuasive benefits of encryption lead many companies to conclude that they should simply encrypt all (or large parts) of their data. The idea behind this is simplification: encrypting everything avoids onerous data classification and risk assessment exercises. However, encryption is not a panacea for good data management. There are several considerations that organizations must make before they deploy encryption.

Should you Encrypt Everything? Downsides to Encryption

Loss of Business Functionality

When data is encrypted using legacy encryption methods, the end result normally looks nothing like the original. This is broadly accepted as a good thing; if an unauthorized individual is looking to decrypt data it is much better not to give them any clues to the original content. But this traditional approach

has a serious downside. Encrypted data cannot be used for normal business processes. Think of any personal data that is used in a business process, such as order taking, customer service, or employee management. Since encrypted data doesn't look like the original it is impossible to process that data. Once a credit card number or date of birth is encrypted it no longer looks like a credit card number or date of birth. Decrypting the data makes it usable, but also exposes the data to risk.

Companies need security, but more importantly they need to function as businesses. Using encryption cannot get in the way of this, and so traditional approaches to encryption are not always suitable for this environment. In fact, the problem is much broader: organizations use copies of personal data for a variety of purposes, such as analytics, reporting, and testing. Encrypted versions of this personal data would render these purposes impossible.

Handoffs Between Encryption Regimes

Most organizations use encryption for very specific purposes. They use it to encrypt data at rest, or in motion, or even while it is in use (being processed). But these use cases are usually dealt with discretely and individually. For example, if encrypted data is required to be transmitted, it is very common to decrypt the data from its at rest state. The data is then re-encrypted before transmitting it across a network, for it to be decrypted when it arrives – this is how SSL/TLS works. If the transmitted data needs to be encrypted when it arrives, then this is done as a secondary action.

These handoffs between separate stages in a data handling process demonstrate the vulnerabilities of using encryption in a fragmented sense. A better solution would be to encrypt the data once and have it stay encrypted as it moves from one location to another across a network.

Granularity of Encryption

Many organizations encrypt data at rest, and this is often regarded as good practice. Notwithstanding the potential risk once the data is moved, data can be regarded as being stored in a secure state. The most common approach to achieving this is to encrypt at file or even disk level. This holistic method is ideal in situations where encryption is built into hardware and tracking application activity is not important, as it takes little overhead to activate native encryption and plug into management systems using interoperability standards. System-level encryption is low-complexity to deploy and automate using enterprise key management approaches for storage/server.

The downside of this approach is that it is an all-or-nothing approach. The whole file, or in the case of full disk encryption (FDE) an entire volume, is encrypted as a whole. But what if more granularity is required? This is particularly true with structured data, where it is not necessary to encrypt at a wholesale level: only certain types of data need to be encrypted. All data is not equal in sensitivity, even within records/rows/documents.

Key Management

The final drawback to encryption, especially encryption of data-in-use and data-in-motion, is that it is not trivial, or rather, key management is not trivial. Many organizations believe that encryption is a relatively straightforward process involving algorithms. In fact, encryption is more about the process of managing keys – get this wrong, and at best you render the encryption pointless, and at worst you suffer a catastrophic loss of data. The consequences of losing keys drive companies to make regular backups of their keys, which of course also need to be secured.

But few industry segments, let alone specific companies, have a strong history of managing encryption keys. Beyond financial services, telecommunications, and the defense sector, most industries have limited experience in using encryption in their business processes. Many organizations avoid encryption due to the complexity of key management.

A compounding factor in key management is the rapid proliferation of keys due to the number of identities supported. This traditionally applies in B2C environments, such as retail and government services, but it will increasingly be an issue in Internet of Things (IoT) environments, where sensors and devices will also require protection through encryption. Coping with the management of keys to protect data as it moves through the enterprise and through networks, at this scale, measured possibly in the millions, presents a massive complexity overhead.

Encryption Works Best When…

Encryption can be a powerful tool to protect data for a variety of reasons, including compliance and IP protection. But it also has downsides that organizations need to understand.

Encryption works best when:

- Simple (but not simplistic) key management is deployed
- Data is encrypted at a granular level, below file or disk encryption, to field-level
- It works across all technology platforms, from legacy mainframes to the Internet of Things
- It works at scale and provides coverage across all data classes

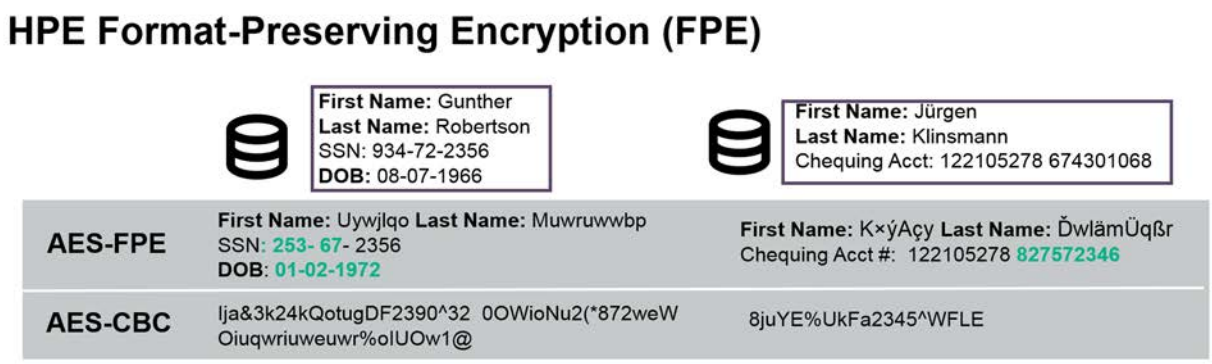
FORMAT-PRESERVING ENCRYPTION AS AN APPROACH TO GDPR COMPLIANCE

Format-preserving encryption is a standards-based approach (based on the NIST SP 800-38G standard) to addressing the requirements of compliance while avoiding the drawbacks of traditional encryption. As the name suggests, data encrypted using FPE retains its characteristics; an encrypted credit card number still looks like a credit card number. FPE can be applied to any structured data formats, such as a telephone number, date of birth, or an email address.

Importantly, FPE preserves the original format of the data. This means that business processing can continue as if the data was unencrypted, without revealing the actual data. It also enables other back-office functions such as backup, analytics, and reporting activities to proceed as normal. The critical advantage here is that it is not necessary to decrypt data in order for it to be useful to the majority of business processes.

FIGURE 1

How Encrypted Data Output Differs Between FPE and Traditional Encryption



Source: Hewlett Packard Enterprise, 2016

In addition, because the data is encrypted with a high degree of granularity, there is limited need to decrypt it as it is being sent around the organization. Thus the handoffs between encryption of data at rest, data in use, and data in transit can be minimized.

FPE also enables field-level encryption, thus providing granularity of encryption as determined by the business process. Encryption is configurable, allowing a variable strength of encryption dependent upon the sensitivity of the data. Granularity also allows selective decryption of specific data fields, without decrypting the whole record, file, or disk.

Finally, FPE also achieves the objectives of pseudonymization, because it protects the data while it is being processed. Essentially, FPE meets both the encryption *and* pseudonymization requirements of GDPR simultaneously.

SIMPLIFYING ENCRYPTION THROUGH STATELESS KEY MANAGEMENT

Stateless key management is an approach that addresses some of the fundamental challenges that organizations face in encrypting their data. It is stateless because keys are generated on demand, based on the identity of the user requesting access. This improves overall scalability of the system and simplifies key management, with no key replication required.

Stateless key management sounds simple. In fact, the complexity is hidden from the enterprise, and so the effect is indeed to simplify key management.

OTHER CHALLENGES IN COMPLYING WITH GDPR

GDPR is not all about security. Much of it regards process, which defines the steps people and technology need to take, and auditing, which checks that these steps have been taken. It's important therefore to make it as hard as possible for people to stray from the defined process (unwittingly or otherwise).

For large volumes of data at rest, in the datacenter for instance, highly scalable infrastructure-level encryption is often desirable, as long as key management is automated as far as possible. This approach to encryption provides a base layer of protection, and is also useful if you don't know what data you need to encrypt, or need to protect the infrastructure as a whole. Infrastructure key management solutions, such as HPE Enterprise Secure Key Manager (ESKM), are a good choice as they provide a centralized key management hardware-based solution for unifying policy and audit visibility, by automating an organization's data-at-rest encryption controls. This is achieved by creating, protecting, serving, and auditing access to encryption keys for secure, reliable administration, and providing extensive interoperability across infrastructure environments via support for the OASIS KMIP standard. HPE ESKM also works in both private and public cloud environments, such as BYOK scenarios where maintaining control of keys on-premise is necessary in co-located, potentially untrusted environments.

Organizations also worry about sending data outside the perimeter via email, a frequent requirement for normal business operations, but also a popular method of data exfiltration by hackers. Most companies therefore use an email security solution, such as HPE SecureMail, that protects personal data sent by email. HPE SecureMail also benefits from the advantages of using stateless key management.

CONCLUSION

Compliance with GDPR is tough and multifaceted. The requirements are broad ranging, cutting across a variety of business functions such as IT, HR, and marketing, as well as risk and compliance. Data security cannot solve all the issues that GDPR raises. Consent, the right to be forgotten, and data portability are just some examples of GDPR requirements that are beyond data security.

However, format-preserving encryption is a useful approach to ensure that data has a core level of protection. GDPR places encryption at the heart of data protection and security, and also understands that encryption is a mitigating factor should a breach occur.

Organizations need to beware, though, of an "encrypt everything" approach. Encryption is not free, in terms of the impact to processing and utility of the data, nor is it trivial. Enterprises need to understand the complexity that accompanies encryption. Importantly, encryption is not a one size fits all technology, and it needs to be done with care. Technologies such as FPE improve the granularity of encryption while preserving business process functionality, and stateless key management helps to minimize the management downsides of encryption.

Organizations considering FPE also need to consider which data to encrypt. It is important, therefore, to discover and then classify personal data in the enterprise. This will help inform a risk assessment of personal data, which drives the decision process regarding which data is encrypted, and how.

GDPR compliance may be tough, but it can be substantially simplified through the use of FPE and stateless key management.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC U.K.

IDC UK
5th Floor, Ealing Cross,
85 Uxbridge Road
London
W5 5TH, United Kingdom
44.208.987.7100
Twitter: @IDC
idc-community.com
www.idc.com