Can EU General Data Protection Regulation Compliance be Achieved When Using Cloud Computing?

Bob Duncan  
Business School  
University of Aberdeen, UK  
Email: bodbuncan@abdn.ac.uk

Abstract—The forthcoming EU General Data Protection Regulation (GDPR) will come into effect across the EU on 25th May 2018. It will certainly be the case that a great many companies will be inadequately prepared for this significant event. While a great many companies who use traditional in-house distributed systems are likely to have a hard enough job trying to comply with this new regulation, but those businesses who use any form of cloud computing face a particularly difficult additional challenge, namely the Cloud Forensic Problem. It is not enough that cloud use presents a far more challenging environment, but that the cloud forensic problem presents a far more difficult barrier to compliance. This problem arises due to the fact that all computing systems are constantly under serious attack, but once an attacker gains a foothold in a cloud system and becomes an intruder, there is very little to prevent the intruder from helping themselves to any manner of data covered by the GDPR, either by viewing it, modifying it, deleting it or ex-filtrating it from the victim system. Worse, there is nothing to prevent the intruder from gaining sufficient privileges to then completely delete all trace of their incursion, possibly deleting far more records than they need to in the process. We address exactly what the requirements of EU GDPR compliance are, consider whether this can be done without resolving the Cloud Forensic Problem, and propose some approaches to mitigate this problem, and possibly the massive potential fines that could then be levied.

Keywords—EU GDPR; Compliance; Cloud computing; cloud forensic problem.

I. INTRODUCTION

The forthcoming EU General Data Protection Regulation (GDPR) [1], is likely to present one of the greatest compliance challenges faced by companies across the globe. Every company that trades anywhere on earth, should they deal with even a single EU resident, must ensure they are compliant with the EU GDPR. If that company suffers a security breach and the records of any EU citizen are compromised, then the jurisdiction of the GDPR will extend globally, and that company may be pursued and fined significant sums of money.

Achieving information security is a big enough challenge for companies who use conventional distributed network systems, but once companies start using cloud systems, the challenge increases exponentially. There are many reasons for this, mostly arising from the complexity of the additional relationships, and agendas, of different participant companies involved in cloud systems. Much research has been carried out to attempt to resolve these problems e.g., [2], [3], [4], [5], [6].

The most challenging, and as yet, unresolved issue is the cloud forensic problem, otherwise known as “The elephant in the room.” Pretty much everyone knows about it, yet nobody is prepared to discuss it, let alone try to resolve the problem, due to the difficulty of the challenge it presents. The forthcoming implementation of the EU GDPR means that heads can no longer be left in the sand. This will not present an acceptable defence.

If any company using cloud is unable to resolve the cloud forensic problem, we suggest this will present such a fundamental issue that it will be impossible for that company to comply with this new regulation. As far back as 2011 and in subsequent years [7], [8], [9], [10], a great deal of research was focussed on trying to resolve this issue, yet it is clear from looking at regulatory fines for breaches that the message is not getting through.

In 2012, Verizon estimated that a total of 174 million data records were compromised [11]. By 2017, this had increased to an estimated 2 billion records lost or compromised in the first half of 2017 alone [12]. Yahoo disclosed a 1 billion compromised account breach in the 2013 attacks, yet when Verizon took over Yahoo last year, it turned out that ALL 3 billion accounts had been compromised [13].

In Section II, we take a look at the implications of non-compliance for any company that falls under the jurisdiction of the forthcoming EU GDPR. In Section III, we identify what the Cloud Forensic Problem is, and address why it is such a challenging problem to overcome. In Section IV, we ask whether it is possible to attain compliance without addressing the cloud forensic problem. In Section V, we address the minimum requirements required necessary to achieve compliance. In Section VI, we look at what achieving the minimum requirements will allow us to do. In Section VII, we consider the limitations of this work, and in Section VIII, we discuss our conclusions.

II. THE EU GENERAL DATA PROTECTION REGULATION

Why should companies be concerned about compliance with the forthcoming EU GDPR [14]? Perhaps the maximum fine for being non-compliant, and suffering a serious cyber breach of the greater of €20 million or 4% of global turnover might serve to grab their attention. We should therefore take a closer look at the detail of the regulation.

The Article 29 Working Party [15] was set up by the European Commission under the terms of Article 29 of the Data Protection Directive in 1996, and its main stated missions are to:
• Provide expert advice to the States regarding data protection;
• Promote the consistent application of the Data Protection Directive in all EU state members, as well as Norway, Liechtenstein and Iceland;
• Give to the Commission an opinion on community laws (first pillar) affecting the right to protection of personal data;
• Make recommendations to the public on matters relating to the protection of persons with regard to the processing of personal data and privacy in the European Community.

During the time it has been active, the Article 29 Working Party has overseen the evolution of the GDPR, and has seen thousands of amendments proposed. One of the best proposals was the requirement to report all breaches “... within 72 hours of the breach occurring”, which would have had the impact of ensuring that all organisations would give security top priority in order to achieve compliance. However, following much lobbying, this was watered down to “... within 72 hours of discovery of a breach.” This rather takes the urgency away from organisations.

On the other hand, another key amendment involved broadening the scope of the regulation, from all organisations anywhere in the EU, to any organisation anywhere in the globe, which stores privately identifiable information relating to any individual resident anywhere in the EU. This will certainly get the attention of far more organisations than would have been the case had it been an EU only requirement.

In the next three subsections, we have a look at how the GDPR seeks to streamline activities for both organisations and data subjects; how the GDPR will use enforcement mechanisms to ensure compliance; and what happens in the event of a data breach.

A. The Streamlining Goals of the GDPR

1) For Organisations: The idea for organisations is to streamline compliance by providing:

- a single set of rules which would apply anywhere in the EU and by using the One Stop Shop approach, covered by Articles 46 to 55 of the GDPR, this would make for a streamlined approach for all organisations, whether based inside or outside the EU.

2) For Data Subjects: The idea for data subjects is to make the whole process for them much simpler by providing:

- Right of Access (under Article 15) - which gives data subjects the right to access their personal data held by any company subject to compliance with the GDPR;
- Right to Erasure (under Article 17) - which gives data subjects the right to have erasure carried out on certain data held by organisations about the data subject on any one of a number of grounds including non-compliance with article 6.1 (lawfulness) that includes a case (f) where the legitimate interests of the controller is overridden by the interests or fundamental rights and freedoms of the data subject;
- Data Portability (under Article 20) - data subjects have certain rights to data portability (particularly in the case of social media accounts), whereby a person shall be able to transfer their personal data from one electronic processing system to and into another, without being prevented from doing so by the data controller;
- Data Protection by Design and by Default (under Article 25) - seeks to ensure that all data subjects can expect privacy by design and by default, that has been designed into the development of business processes for products and services. This requires that privacy settings must be set at a high level by default and that technical and procedural measures should be taken care of by the controller in order to make sure that the processing, throughout the whole processing lifecycle, complies with the regulation. A report by the European Union Agency for Network and Information Security (ENISA) [16], elaborates on what needs to be done to achieve privacy and data protection by design. It specifies that encryption and decryption operations must be carried out locally, not by remote service, because both keys and data must remain in the power of the data owner if any privacy is to be achieved. Furthermore, it specifies that outsourced data storage on remote clouds is practical and relatively safe, as long as only the data owner, not the cloud service, holds the decryption keys;
- Consent by Data Subjects - data subjects must have given their consent for data about them to be processed, thus providing a lawful basis for processing.

3) A Lawful Basis for Processing: The data subject must have given consent which must be explicit for data collected and the purposes data is used for (Article 7; defined in Article 4). Data controllers must be able to prove “consent” (opt-in) and consent may be withdrawn. Consent for children must be given by the child’s parent or custodian, and must be verifiable (Article 8). Such consent to the processing of his, her or their personal data for one or more specific processing purposes, must be:

- necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
- necessary for compliance with a legal obligation to which the controller is subject;
- necessary in order to protect the vital interests of the data subject or of another natural person;
- necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
- necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

B. Enforcement Mechanisms

- Appointing a Data Protection Officer - this person would be required for all data processor organisations,
and a person with expert knowledge of data protection law and practices should assist the controller or processor to monitor internal compliance with this Regulation. The appointment of a DPO within a large organization will be a challenge for the Board as well as for the individual concerned, due to the myriad governance and human factor issues that organisations and companies will need to address given the scope and nature of the appointment. In addition, the post holder will need to create their own support team and will also be responsible for their own continuing professional development as they need to be independent of the organization that employs them, effectively as a "mini-regulator";

- Ensuring Compliance with the GDPR, by checking that all the correct mechanisms are properly defined and in place, mainly through compliance demonstration, e.g. the data controller should implement measures which meet the principles of data protection by design and data protection by default. Such measures include the process of pseudonymising (Recital 78), i.e., by means of encryption, which process should be completed as soon as is practically possible.

- The GDPR seeks to provide Responsibility and Accountability by all parties involved in data processing, with expanded notice requirements covering retention time for personal data, and contact information for data controller and data protection officer. Automated decision-making for individuals, including algorithmic means of profiling (Article 22), which is regarded as contestable, similar to the Data Protection Directive (Article 15), receive particular attention. The expectation is that all actors involved in the whole process of data processing will behave responsibly and will be fully accountable for their actions. Data Protection Impact Assessments (Article 35) have to be conducted when specific risks occur to the rights and freedoms of data subjects. Risk assessment and mitigation is required and prior approval of the Data Protection Authorities (DPA) is required for high risks. Data Protection Officers (Articles 3739) are to ensure compliance within organizations.

C. In the event of a Data Breach

In the event of a data breach, under the GDPR, the Data Controller will be under a legal obligation to notify the Supervisory Authority without undue delay. The reporting of a data breach is not subject to any de minimis standard and must be reported to the Supervisory Authority within 72 hours after having become aware of the data breach (Article 33). Individuals have to be notified if adverse impact is determined (under Article 34), unless the data was encrypted. In addition, the data processor will have to notify the controller without undue delay after becoming aware of a personal data breach (under Article 33).

1) Sanctions: The following sanctions can be imposed:

- a warning in writing in cases of first and non-intentional non-compliance;
- regular periodic data protection audits;
- a fine of up to €10 million or up to 2% of the annual worldwide turnover of the preceding financial year in case of an enterprise, whichever is greater, where there has been an infringement of the following provisions (Article 83, Paragraph 4[18]):
  - the obligations of the controller and the processor pursuant to Articles 8, 11, 25 to 39 and 42 and 43;
  - the obligations of the certification body pursuant to Articles 42 and 43;
  - the obligations of the monitoring body pursuant to Article 41(4).
- a fine up to €20 million or up to 4% of the annual worldwide turnover of the preceding financial year in case of an enterprise, whichever is greater, where there has been an infringement of the following provisions: (Article 83, Paragraph 5 & 6[18]):
  - the basic principles for processing, including conditions for consent, pursuant to Articles 5, 6, 7 and 9;
  - the data subjects’ rights pursuant to Articles 12 to 22;
  - the transfers of personal data to a recipient in a third country or an international organisation pursuant to Articles 44 to 49;
  - any obligations pursuant to Member State law adopted under Chapter IX;
  - non-compliance with an order or a temporary or definitive limitation on processing or the suspension of data flows by the supervisory authority pursuant to Article 58(2) or failure to provide access in violation of Article 58(1).

The above details provide the essence of what we need to know in order to understand what information will be required to be delivered in the event of breach, in order for the data processor to be compliant with the GDPR. In the next section, we will take a look at the Cloud Forensic Problem, and why it is such a difficult problem, not only from the security perspective, but also from the GDPR compliance problem.

III. THE CLOUD FORENSIC PROBLEM (AND WHY IT IS SUCH A DIFFICULT PROBLEM)

As we have already stated, all computing systems are constantly under serious attack, but once an attacker gains a foothold in a cloud system and becomes an intruder, there is little to prevent the intruder from helping themselves to any amount of data covered by the GDPR, either by viewing it, modifying it, deleting it or ex-filtrating it from the victim system [17], [18], [19]. Worse, there is nothing to prevent the intruder from gaining sufficient privileges to then completely delete all trace of their incursion, possibly deleting far more records than they need to in the process, leading to further problems for business continuity.

This is often known as “The elephant in the room” in cloud circles. Pretty much everyone knows about it, yet nobody is prepared to discuss it, let alone try to resolve the problem, due to the difficulty of the challenge it presents. Make no mistake, this is a serious challenge to defend against, let alone overcome. However, not only is it a serious challenge for organisations using cloud, it also presents a major obstacle to compliance with the GDPR.
Once all trace of the intrusion has been deleted, there will be very little forensic trail left to follow, meaning many companies will be totally unaware that the intrusion has taken place, let alone understand what records have been accessed, modified, deleted or stolen. All too often, companies will believe they have retained a full forensic trail in their running instance, but often forget that without special measures being taken to save these records off-site [2], they will vanish when the instance is shut down.

Currently, in any cloud based system, there must be a complete and intact audit trail in order for the breached organisation to be able to tell which records have been accessed, modified, deleted or stolen. Where the audit trail and all forensic records have been deleted, there remains no physical means for any organisation to be able to tell which records have been accessed, modified, deleted or stolen, putting these organisations immediately in multiple breaches of the GDPR.

IV. IS IT POSSIBLE TO ACHIEVE COMPLIANCE WITH THE EU GDPR WITHOUT ADDRESSING THE CLOUD FORENSIC PROBLEM?

The short answer is, of course, it is not! For the reasons outlined in the previous section, we can see that there is indeed nothing to prevent an intruder from destroying every scrap of forensic proof of their incursion into any current cloud system. It is clear that any form of forensic record or audit trail can not therefore be safely stored on any running cloud instance.

This means that the only safe method of storage of forensic data will be somewhere off-site from the running cloud instances. Clearly, the off-site storage must be highly secure, preferably stored in an append-only database, and should especially be held in encrypted format, with all encryption keys held elsewhere.

Doubtless some will say that as long as they are not breached, then they will not be in breach of the GDPR. While that may very well be true, how will they be able to tell whether they have not been breached, against the circumstance where they have been breached, and the breach has been very well covered up. They will have no means of knowing, let alone proving the point.

Let us suppose that a complaint is made to the regulator, the organisation will have no means of proving that the data has not been tampered with. Equally, if the breach has been extremely well covered up, they will neither have the means of complying with the requirement to: a) report the breach within 72 hours, nor b) have any means of determining which records have been accessed, modified, deleted or stolen. Let us now suppose that the conversion of private data has yet to be encrypted, and worse, that the encryption and decryption keys are held on the cloud instance “for convenience”. If we were to receive a request from any users whose data had just been compromised, we would be unable to comply with the request, meaning we would now be looking at multiple breaches, thus causing the fine level to escalate to the higher level, as outlined in Subsubsection II-C1.

V. THE MINIMUM REQUIREMENTS TO ACHIEVE COMPLIANCE WITH THE GDPR

We have seen that to do nothing would not be a viable option as far as GDPR compliance is concerned. Attacks will continue unabated. We must therefore be prepared and armed with whatever tools we can develop to ensure we achieve as high a level of compliance as we possibly can. We therefore need to consider what the absolute minimum technical requirement might be to attain our objective of GDPR compliance. We know that under the GDPR the organisation must be able to:

- provide a Right of Access (under Article 15) to personal data by data subject, if requested;
- provide the means to comply with a Right to Erasure (under Article 17) by data subject, subject to the appropriate grounds being met;
- provide privacy by design;
- in the event of a data breach, report the breach to the Supervisory Authority within 72 hours after having become aware of the data breach (Article 33). The breach must also be reported to the controller without undue delay after becoming aware of a personal data breach;
- in the event of a data breach, notify the data subject if adverse impact is determined (under Article 34), unless the data was encrypted;

In the case of the first requirement, we would require to ensure the provenance and veracity of the contents of the database. In the case of the second requirement, if appropriate, the same provision would apply.

In the case of the third requirement, the cloud system must be designed in accordance with the recommendations of the Article 29 Working Party [20], which suggests the reports produced by ENISA should be followed. This report [21] specifies that encryption and decryption operations must be carried out locally, not by remote service, because both keys and data must remain in the power of the data owner if any privacy is to be achieved. Furthermore, it specifies that outsourced data storage on remote clouds is practical and relatively safe, as long as only the data owner, not the cloud service, holds the decryption keys. ENISA have also produced a stream of other relevant reports, including a Cloud Risk report in 2009 [22], and recommendations for certification in 2017 [23].

In the case of the fourth requirement, we would require to ensure the provenance and veracity of the contents of the database. In the case of the fifth requirement, where the data is not yet encrypted, the same provision would also apply. However, it should be stressed that it will always be preferable to ensure data is encrypted before it leaves the control of the data owner.

It is clear that where no steps have been taken to ensure the cloud forensic problem has been mitigated, the organisation will fail on every count. Thus, as a minimum, we need to ensure the following steps are taken:

- all personal data should be encrypted, and this should be performed locally;
- the encryption and decryption keys should not be maintained on the cloud instance;
- a full audit trail of the entire database must be maintained off-site;
• full forensic records of all users having accessed the database and carried out any commands on the database must be collected and stored off-site.

VI. WHAT WILL THE MINIMUM REQUIREMENTS ALLOW US TO DO?

Let us now assume that we have completed the minimum requirements. Can we now be sure that we can be compliant with the provisions of the GDPR? We must therefore look at each of the five reporting requirements in turn to establish whether we will be able to meet these requirements.

1) First, if a data subject serves us with a Right of Access request, can we respond in the affirmative? We are now sure that we hold the subject’s data securely, in encrypted format in our database. Further we can prove that the data has only been accessed by duly authorised persons, and that the data records have neither been modified, stolen nor deleted. We are therefore compliant on the first requirement;

2) Next, if a data subject serves us with a right to Erasure notice, can we comply with that request. Assuming the request can be legitimately carried out and is not prohibited by statute, then since we can correctly identify the private data held about the data subject, then there is no reason why we would be unable to delete the appropriate data as requested. Accordingly, we would be compliant on the second requirement;

3) Next, can we provide privacy by design? Since we can comply with the first two requirements, this is a clear indication that we are potentially capable of supplying privacy by design;

4) In the event of a data breach, can we report the breach to the Supervisory Authority within 72 hours of discovery? In the case of a data breach, we will not only be able to notify the breach within 72 hours of discovery, we will actually be able to notify within 72 hours of the occurrence of the breach. In addition, since we will retain full forensic data and audit trails for the system, we will also be able to provide very precise details of which records were accessed and read, which might have been modified, with full details of what modifications were made, which records were deleted, and which records were exfiltrated from the system. Not only that, but we will be able to provide full details of how the perpetrators got into the system and where they forwarded any stolen records, which means we can identify precisely which records were compromised, thus ensuring we would be beyond fully compliant;

5) In the event of a data breach, would we be able to notify the data subject if adverse impact is determined (under Article 34)? In the event of a data breach, we would be able to identify every single record attacked, and identify every single data subject affected. Since the full records would already be encrypted, we would not be required to notify the data subjects, but would be fully capable of so doing. This would mean we would again be beyond fully compliant.

Thus, we can reasonably claim that we would be in a position to be fully compliant with all the requirements of the GDPR, thus providing an exceptionally high level of privacy on behalf of all data subjects. Thus, the level of exposure of data subjects would be extremely minimised, thus ensuring compliance with the regulation, and therefore the likelihood that we would be able to fully mitigate any penalty that would otherwise be applied by the regulator.

Contrast this position with the case where cloud users do not take these mitigatory steps. In every requirement - they would be non-compliant, thus exposing the enterprise to the full extent of penalties allowed, namely the greater of €20 million or 4% of global turnover.

VII. LIMITATIONS AND DISCUSSION

There are two very important tasks that must be performed in order not to limit the effectiveness of this approach. Since persistent storage in the cloud instance cannot retain data beyond its currently running lifetime [2], we must also make sure that all necessary logs and data is stored securely elsewhere. And as the default settings for virtually all database software involves logging being turned off [17], we must ensure this function is turned on in all running cloud instances, again, with the data being stored securely elsewhere.

This prompts the question of what data we require to keep. In order to meet our regulatory compliance requirement, we need to understand the 5 W’s — namely: Who is accessing our system? Where have they come from? What are they looking for? When is this happening? From this data, we should be able to infer the Why? Are they authorised to be in the system, to enter the system the way they have, to look at the data they are trying to access, and at the time they are trying to access it? Deducing the Why can give an indicator of anomalous behaviour.

Many database software offers additional full audit trial capabilities. Each additional capability will require more and more storage resources. A balance will need to be found between the minimum requirement consistent with maintaining performance and a cost effective level of storage. The risk in not utilising all that is on offer, would be that this might compromise security, reducing the ability of achieve compliance.

However, it is clear that a sensible precaution to mitigate this risk would be to encrypt all the data being held on all databases maintained within the system, ensuring that encryption/decryption keys are not stored on the cloud instances. While encryption is not mandatory, in the event of a breach where encryption is not used, the fine levied by the regulator is likely to be much higher as a consequence.

VIII. CONCLUSION

The forthcoming GDPR will certainly present a serious wake up call to a great many companies operating around the globe if they find themselves falling under the jurisdiction of this new regulation. In this paper, we have considered whether it is possible to achieve regulatory compliance where any organisation is using cloud computing. Again, we reiterate that without suitable precautions being put in place, the answer is a resounding “No!”.

We have outlined the key requirements from the regulation to which all organisations falling under its jurisdiction must comply. We have identified the currently unresolved “Cloud Forensic Problem” as presenting the largest obstacle to achieving compliance.
We have proposed how this challenging problem may be approached to ensure that cloud users can be fully compliant with this new regulation, with little more than being sensibly organised. Clearly, additional cost will require to be incurred, and there may be a small impact on latency, but these costs could significantly mitigate the possibility of a huge regulatory fine in the event of a breach. It is also likely that this approach will ensure faster discovery of the occurrence of a breach, thus minimising the potential impact on business continuity.

REFERENCES


