But my emails.

Contributors **1**

Abstract **1**

A brief history of the Schrems II ruling **2**

Implications for Software-as-a-Service **2**

Email use case **4**

Functions 4

SaaS solutions 4

Email encryption 5

Proposed architecture **6**

General 6

Functional aspects 6

Regulatory aspects 7

Challenges 13

Conclusion **13**

Contributors

Adam Leon Smith FBCS CITP, CTO Dragonfly.
Mark Potkewitz
David Clarke

Abstract

This paper summarises the output of a collaborative open webinar held between technical and regulatory experts at the Open Security Summit and the British Computer Society in November 2020, immediately following guidance from the EPDB with recommendations for compliance following the Schrems II ruling.

It represents the discussion of experts, follow-up analysis and elaboration. It is not formally endorsed by any organisation, but aims to progress the discussion around potential ways to comply with the regulator recommendations. Through the exploration of a very common use case (email), it proposes an example architecture for Software-as-a-Service (SaaS) in light of Schrems II EPDB guidance

A brief history of the Schrems II ruling

In June of 2013, journalist Glenn Greenwald[1] published revelations[2] from Edward Snowden,[3] a Booz-Allen[4] consultant working for the U.S. National Security Agency (NSA).[5] Amongst the documents identified as those taken by Snowden indicated the widespread,[6] court-authorised,[7] warrantless collection[8] of telephony[9] and Internet metadata and browsing data[10] by the United States Intelligence Community (USIC)[11] of callers inside the United States, under Section 215[12] of the USA Patriot Act,[13] and outside the United States, under Section 702[14] of the Foreign Intelligence Surveillance Act[15] (FISA) Amendments Act of 2008[16] (FISA Amendments Act).

Following the exposure of widespread surveillance programmes like PRISM[17] and X-Keyscore,[18] Austrian lawyer and privacy activist, Maximillian Schrems, TBD… <Mark>

Implications for Software-as-a-Service

Any ability to remotely access the data, or retrieve it, means that the data is being transferred to the country in which it is being accessed.

This does not mean the processing is unlawful. However, if the country in which the data is processed or accessed is not in the European Union, and is not considered an adequate country, a number of additional steps should be taken[19]. It should be noted that the US is not considered an adequate country[20], and the UK is not expected [ref?] to gain an adequacy decision after the transition period[21].

---

[1] https://twitter.com/ggreenwald
[2] https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order
[3] https://www.wired.com/2014/08/edward-snowden/
[4] https://www.boozallen.com/
[5] https://www.nsa.gov/
[6] https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order
[7] https://www.theguardian.com/commentisfree/2013/jun/19/fisa-court-oversight-process-secrecy
[8] https://www.theguardian.com/world/2013/jun/20/fisa-court-nsa-without-warrant
[9] https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order
[10] https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data
[11] https://www.dni.gov/index.php/what-we-do/members-of-the-ic
[12] https://www.law.cornell.edu/uscode/text/50/1861
[13] https://www.congress.gov/bill/107th-congress/house-bill/3162
[14] https://www.law.cornell.edu/uscode/text/50/1881a
[15] https://www.govinfo.gov/content/pkg/STATUTE-92/pdf/STATUTE-92-Pg1783.pdf#page=15
[16] https://www.congress.gov/bill/110th-congress/house-bill/6304
[17] https://www.theguardian.com/world/2013/jun/08/nsa-surveillance-prism-obama-live
[18] https://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data
[19] https://edpb.europa.eu/sites/edpb/files/consultation/edpb_recommendations_202001_supplementarymeasurestransferstools_en.pdf
[20] https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en
[21] https://www.ianbrown.tech/2020/11/30/the-uks-intelligence-activities-and-gdpr-inadequacy/

The following steps[22] must be taken to access the data from a country that is not deemed adequate:

Firstly, if the legal entity that controls the data in the user country is not the legal entity that processes (and or controls) the data in the other country, then the entities must establish appropriate safeguards in the form of contracts between them. Standard clauses are available from the EU. This is relatively straightforward.

Secondly, you must assess whether the contract is enforceable. It is not considered enforceable in the US. As the public authorities are not a party to the contract, they can force a US entity or person to access the data, even if this is remote access. There is no precedent to determine whether UK law will be considered sufficient in this regard.

Finally, if it is not considered an enforceable contract (e.g. the US), then additional supplementary measures must be taken. Contractual and organisational measures are not sufficient.

The EPDB give several recommended scenarios that would be effective:

- Data is transferred in encrypted form and decryption is not possible by a person in the other country.
- Data is pseudonymised, and similarly, de-pseudonymisation is not possible in the other country.
- Data is only transiting the third country in encrypted form.

It specifically gives the following examples that cannot be effective in achieving sufficient protection:

A data exporter uses a cloud service provider or other processor to have personal data processed according to its instructions in a third country. If:
1. a controller transfers data to a cloud service provider or other processor,
2. the cloud service provider or other processor needs access to the data in the clear in order to execute the task assigned, and
3. the power granted to public authorities of the recipient country to access the transferred data goes beyond what is necessary and proportionate in a democratic society,

22 https://edpb.europa.eu/sites/edpb/files/consultation/edpb_recommendations_202001_supplementarymeasurestransferstools_en.pdf

Email use case

Functions

The basic function of an email service comprises the routing and delivery of text or multimedia messages between one or more recipients at one or more entities. If this was the only function, this would be an easier problem to solve. Typically, SaaS solutions provide:

- Client-side software for accessing and managing emails;
- Web-based access for accessing and managing emails;
- Cloud-based search of email content;
- An email address lookup function for other individuals in the same organisation;
- An email address book storage function for external individuals;
- Features for managing compliance and data loss prevention, such as scanning emails for particular keywords or data types and for example, routing them for approval;
- Security features such as spam and malware detection;
- Backup.

SaaS solutions

Email SaaS is a duopoly, with Google's G-Suite and Microsoft's Office 365 controlling almost the entire market[23]. Whilst many large companies will use on-premise email servers, small and medium-size enterprises are more likely[24] to use SaaS versions for ease-of-use and a reduction in complexity and setup costs.

Email represents a huge risk in terms of personal data. Users rarely delete email, and the minimum mailbox size for Office 365 is 50 gigabytes[25]. To explore the most obvious example, employers when hiring generally process CVs, identification documents, bank details, and social security information. Some of this might be sensitive data, for example, data pertaining to ethnicity.

In the case of both Microsoft and Google SaaS, it is hard to see how their use can be compliant with GDPR. Google says[26] the following:

[23] https://www.datanyze.com/market-share/office-suites--370/office-365-vs-g-suite
[24] https://enlyft.com/tech/products/microsoft-office-365
[25] https://docs.microsoft.com/en-us/office365/servicedescriptions/exchange-online-service-description/exchange-online-limits#mailbox-storage-limits
[26] https://cloud.google.com/blog/products/identity-security/google-clouds-commitment-to-eu-international-data-transfers-and-the-cjeu-ruling

"Given the CJEU has upheld the MCCs, it is important to know that your use of G Suite and Google Cloud Platform meets GDPR's standards for transfer of personal data outside of the EU."

However, this is clearly a false premise. The fact that the CJEU has upheld the legitimacy of contracts as a compliance tool, does not lead to automatic compliance when they are used. This is clear from the latest guidance.

Microsoft[27] makes a similar argument "*Although today's ruling invalidated the use of Privacy Shield moving forward, the SCCs remain valid. Our customers are already protected under SCCs."*

We can contrast these statements with the latest guidance from the EPDB which states *"Contractual and organisational measures alone will generally not overcome access to personal data by public authorities of the third country".*

Email encryption

End to end encryption of emails is a well-understood problem that was solved in 1991 with PGP[28]. However, it's uptake has been extremely low. One reason for this is that it typically requires some work to configure that puts off lay users, for example, Microsoft's popular email client Outlook does not support PGP natively, nor does Gmail. Enterprises have generally not favoured end-to-end encryption of email, as they have to balance data security against other compliance requirements. Generally, employers want to be able to access employee's email in a manual or automated fashion for a variety of reasons. Governments also do not encourage the use of end-to-end encryption, ostensibly for public safety reasons[29].

Analysing the generic functions listed above for SaaS email, end-to-end encryption would effectively disable web-based access, cloud-based search, compliance and data-loss prevention, spam and malware detection. An example of this is ProtonMail, which offers full end-to-end encryption, but only meta-data such as subject lines, dates and senders are searchable[30]. Another example is Tutanota[31], that allows for the search of encrypted data. This, open-source work provides an example of how encrypted data can be searched without decryption.

[27] https://blogs.microsoft.com/eupolicy/2020/07/16/assuring-customers-about-cross-border-data-flows/
[28] https://en.wikipedia.org/wiki/Pretty_Good_Privacy
[29] https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety
[30] https://protonmail.com/support/knowledge-base/search/
[31] https://tutanota.com/blog/posts/first-search-encrypted-data/

Proposed architecture

General

The architecture of software systems is often described using multiple views[32]. These might be logical, physical, or be a view from a business or development process perspective. To build a compliant architecture, it might be useful to introduce a trust view. The trust view should show two layers, a trusted layer, and an untrusted layer.

The trusted layer would need to be hosted in an adequate country, by an organisation that could not be subject to United States law. The untrusted layer could be hosted anywhere, and controlled by anyone. In order to be compliant, the untrusted layer could only process data that was encrypted (or pseudonymised). The data would need to be encrypted using suitable strong algorithms, and the private keys might be held by both the end-user, and the trusted layer of the architecture.

The trusted layer should include private keys, to abstract technical problems from the users, but also to allow for conditional access by the authorities of the country hosting the trusted layer.

Functional aspects

It is assumed that the end-user environment could remain largely the same as it does now, comprising email client applications, and a web browser that could access a web server.

The trusted environment would have a key store, holding asymmetric public and private keys for users, and fetching public keys from a certificate authority for correspondents. It would need to have some kind of web-server, providing user access to email through their browsers.

Most importantly, it would provide encryption and decryption of email content so that data in clear form does not pass to the untrusted layer. This content would, of course, need to be decryptable by the recipients, so PGP architecture and certificate authorities could be used in this layer, abstracting the complexity of this from the end-user.

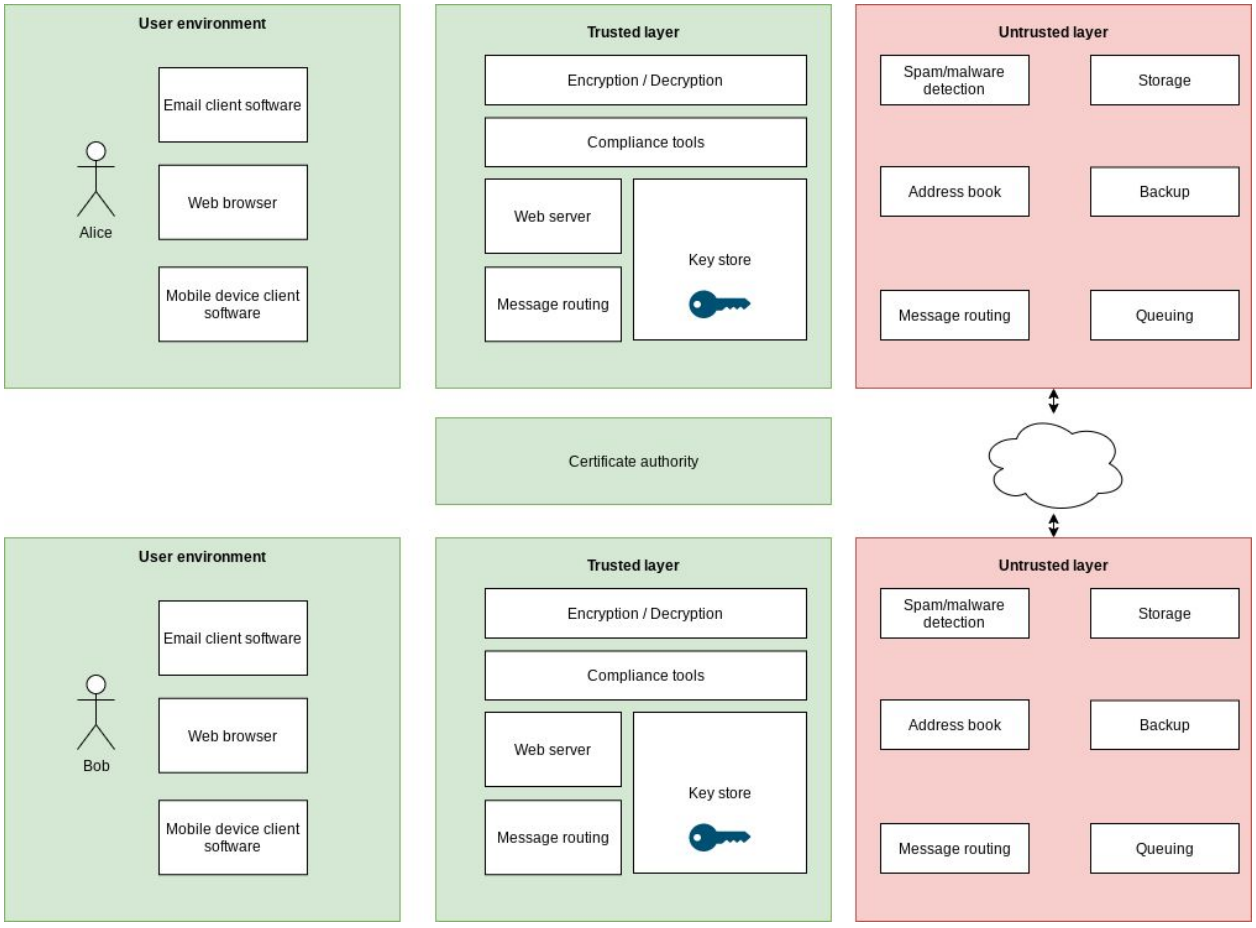
Searching of emails can be processed by the untrusted layer. This is because as long as the search terms are encrypted using the same method as the original email, the encrypted term would match the encrypted content. This is a form of homomorphic encryption[33]. In order to avoid the search term being encrypted with the keys of all correspondents, it might be necessary to store a version of all incoming emails re-encrypted with the user's own key, rather than that of the original sender. The same principle can be applied to address book lookup and storage.

[32] https://en.wikipedia.org/wiki/4%2B1_architectural_view_model
[33] https://en.wikipedia.org/wiki/Homomorphic_encryption

Of greater complexity might be the solutions for spam and malware. These typically use Bayesian probability or machine learning to identify target messages. Machine learning can operate on encrypted data, as long as the encryption keys are constant. Whilst this would provide a solution for algorithms that learn based on data from the same recipient organisation, it may not be possible for centrally developed algorithms (by the untrusted party) to be used across organisations. A possible solution to this might be federated learning.

Compliance functionality and data-loss prevention require the comparison of content against particular rules, for example, regular expressions. This is more challenging to perform on encrypted data than simple matching. It would require some form of homomorphic encryption[34] in order to be performed in the untrusted layer, and as these techniques are in their relative infancy in terms of practical use, it is placed in the trusted layer in this example.

Backup, and other infrastructural elements of the solution (for example alerting and monitoring) can operate as they currently do as they do not require access to data in the clear.



Explain the open-source nature of the trusted environment
It may be that the trusted layer can be an open-source framework or application set.

[34] https://en.wikipedia.org/wiki/Homomorphic_encryption

Regulatory aspects

How the BCS SG addresses the "Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data" summary as follows from EDPB Guidelines.

1. *Know your transfers*
2. *Verify the transfer tool your transfer relies on*
3. *Assess if there is anything in the law or practice of the third country*
4. *To identify and adopt supplementary measures*
5. *Take any formal procedural steps the adoption of your supplementary measure may require*
6. *Re-evaluate at appropriate intervals*

Based on the email use case of the major global email providers (which include a huge amount of extra functionality such as video conferencing, storage etc.that is not in scope for this analysis), the EPDB guidance can be applied as follows:

Step	EDPB Recommendations	Action by Corporate Entity	FISA[35]	Transfer to 3rd country? [36]	Mitigation
1.	Know your transfers	Run Data Discovery Mapping		Pass	BAU
2.	Verify the transfer tool your transfer relies on	Ensure Adequacy Decision		NO	Step 2 assessment has failed

[35] US data importers that fall under 50 USC § 1881a (FISA 702), may extend to cryptographic keys

[36] Can Data be Transferred/Exported to 3rd country or 3rd country corporate actors access it?

3.	Assess if there is anything in the law or practice of the third country	88. A data exporter uses a cloud service provider or other processor to have personal data processed according to its instructions in a third country." the power granted to public authorities of the recipient country to access the transferred data goes beyond what is necessary and proportionate in a democratic society," **Not Possible**	Yes	NO	Yes FISA 702 , does not meet EDPB guidance.
4.	To identify and adopt supplementary measures "to the extent that it addresses the specific deficiencies identified in your assessment of the legal situation in the third country"	79. Encryption keys , under exporter control. 80. Pseudonymisation if original data is not hosted by 3rd country company 85. If data importer is not under FISA 701. 86. the algorithm used for the shared computation is secure against active adversaries **Not Possible**	Yes Importers have access to Keys	NO	Step 3 assessment has failed

5.	Take any formal procedural steps the adoption of your supplementary measure may require	**Not Possible**	Supplementary measures not possible	NO	Step 3 assessment has failed
6.	Re-evaluate at appropriate intervals the level of protection afforded to the data you transfer to third countries and to monitor if there have been or there will be any developments that may affect it.	**Possible**		YES	

Technical and Organisational Measures that will suffice for EDPB Step3 "Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data Adopted on 10 November 2020"

Item	**Issue to mitigate public authorities of the recipient country to access the transferred data goes beyond what is necessary" Criteria**	**Action**	**Meets EDPB guidelines**
1	the personal data is processed using strong encryption before transmission	Importer uses in-country platform with encryption keys owned and Managed In Country. With relevant skills, controls and ability to manage in country of data exporter .	YES

2	the encryption algorithm and its parameterization (e.g., key length, operating mode, if applicable) conform to the state-of-the-art and can be considered robust against cryptanalysis performed by the public authorities in the recipient country taking into account the resources and technical capabilities (e.g., computing power for brute-force attacks) available to them	Importer uses in country encryption algorithm platform with encryption keys owned and Managed In Country. With relevant skills, controls, resources and ability to manage in country of data exporter.	YES
3	. the strength of the encryption takes into account the specific time period during which the confidentiality of the encrypted personal data must be preserved,	Using In country of data exporter developed open source encryption Software	YES
4	the encryption algorithm is flawlessly implemented by properly maintained software the conformity of which to the specification of the algorithm chosen has been verified, e.g., by certification,	Importer uses in country encryption algorithm platform with encryption keys owned and Managed In Country. With relevant skills, controls, resources, certification and ability to manage in country and decrypt only in country.	YES
5	the keys are reliably managed (generated, administered, stored, if relevant, linked to the identity of an intended recipient, and revoked), and	In country encryption algorithm platform with encryption keys owned and managed In country. With relevant skills, controls, resources and ability to manage in country.	YES
6	the keys are retained solely under the control of the data exporter, or other entities entrusted with this task which reside in the EEA or a third country, territory or one or more specified sectors within a third country, or at an international organisation for which the Commission has established in accordance with Article 45 GDPR that an adequate level of protection is ensured	In country encryption algorithm platform with encryption keys owned and Managed In Country. With relevant skills, controls, resources and ability to manage in country., with ownership of encryption keys and access with the data exporter.	YES

In Summary the Technical Requirements to meet EDPB guidelines for the above use case can be met.

The BCS SG Octad for Data Protection

1. Open source software (independent of company and country)
2. Hosted in country of data exporter
3. Managed by a company in the country of the data exporter
4. Ability to decrypt in country of data exporter with open source software
5. Encryption keys owned and managed in country of data exporter
6. Using in country of data exporter where encryption software is developed.
7. Managed in country with appropriately skilled, resourced and certified platform
8. Data exporter retains ownership of Keys

Challenges

This proposed architecture is illustrative, to show how the email use case could be compliant. However practically there are some challenges:

1. This architectural paradigm is conceptual and may present practical challenges in implementation.

2. The impact upon the business model of SaaS providers may prove insurmountable.
3. The big-tech providers have significant cyber-security expertise, the fragmentation of the trusted layer may provide a larger attack surface, and this could be taken advantage of by bad actors if not carefully managed.
4. The trusted layer may become a single point of failure.

Conclusion

Summarise...