SHARDSECURE

Enabling GDPR/Schrems II Compliance

Data protection for cross-border data transfers

Hans-Peter Erlingsson, CEO, Lex Legem Advisory and Consulting
Jesper Tohmo, CTO and Co-Founder, ShardSecure

Introduction

The ruling of the European Union Court of Justice in the Schrems II case has made the adoption of global cloud services by EU/EEA organizations significantly more complicated. This whitepaper demonstrates how ShardSecure's patent-pending approach to data protection helps organizations satisfy specific requirements for cross-border data transfers and GDPR-compliant processing of personal data in cloud services as recommended by the European Data Protection Board (EDPB).

Schrems II ruling and its implications for cloud adoption

On 16 July 2020, the Court of Justice of the European Union ("the Court") invalidated the EU-US Privacy Shield adequacy decision in its Case C-311/18 Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems (called the "Schrems II case"). The Court also cast doubt over the extent to which data transfers from the EU and the EEA to the US (and to other jurisdictions with significantly less data protection) could be legitimized by the European Commission's Standard Contractual Clauses (SCC). The Court upheld the general validity of SCCs as a transfer tool but stressed that SCCs must be supplemented by technical or organizational safeguard measures if the laws or practices of the third country impinged on the effectiveness of those SCCs. The implications of the Schrems II ruling for cloud adoption were groundbreaking. The Court's repeal of the EU-US Privacy Shield adequacy invalidated the entire legal basis for free data flows to the US. Additionally, SCCs and other GDPR-approved transfer tools could no longer be trusted without the implementation of supplementary measures. Overnight, many organizations found themselves in limbo, struggling to balance operational efficiency and information security with privacy and regulatory compliance.

Split or multi-party processing – an EDPB recommended supplementary measure (Use Case 5)

On 18 June 2021, the European Data Protection Board (EDPB) adopted recommendations on the measures that supplement transfer tools to ensure compliance with EU-level data protection. These recommendations clarify the duties of data exporters in international data transfers and expand on which measures do and do not suffice to supplement the transfer tools, including the SCCs. In this context, the EDPB in Use Case 5 explicitly refers to "split or multi-party processing" as a generally acceptable supplementary measure. Splitting information into smaller pieces prior to transmission and distributing those pieces across multiple processors, locations, and jurisdictions in such a way that no piece can be reconstructed by a single processor will effectively eliminate the risk of the laws and practices of third countries impinging on the safeguards of the SCCs and other approved transfer tools.

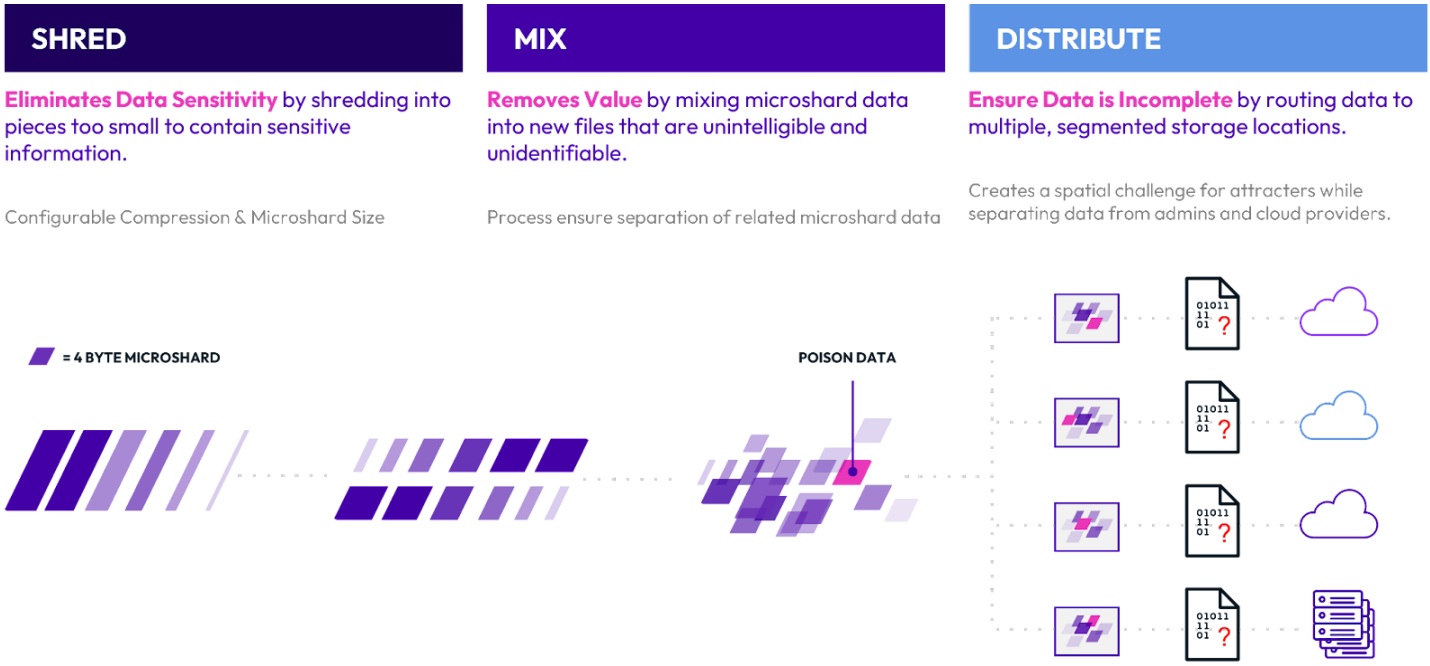
Microshard Technology: A Brief Overview

ShardSecure's Microshard™ technology helps ensure privacy, security, and resilience for data at rest in hybrid- and multi-cloud environments. The microsharding process, described in detail below, helps to ensure that microsharded data is unintelligible and of no value to any unauthorized user.

The solution is a software-based virtual cluster that customers may operate on-premises and/or in the cloud. The solution functions as an abstraction layer between a customer's application servers and a customer's hybrid-cloud or multi-cloud storage. At no time is customer data stored in or read by the solution.

The frontend appears as simple cloud storage via an API and as network storage via an iSCSI module. Applications simply save data to storage as usual, but the data is passed through the microsharding engine , described below, before being stored. The backend distributes the microsharded data to multiple customer-owned storage locations.

The microsharding process itself consists of three main steps: Shred. Mix. Distribute.



Shred

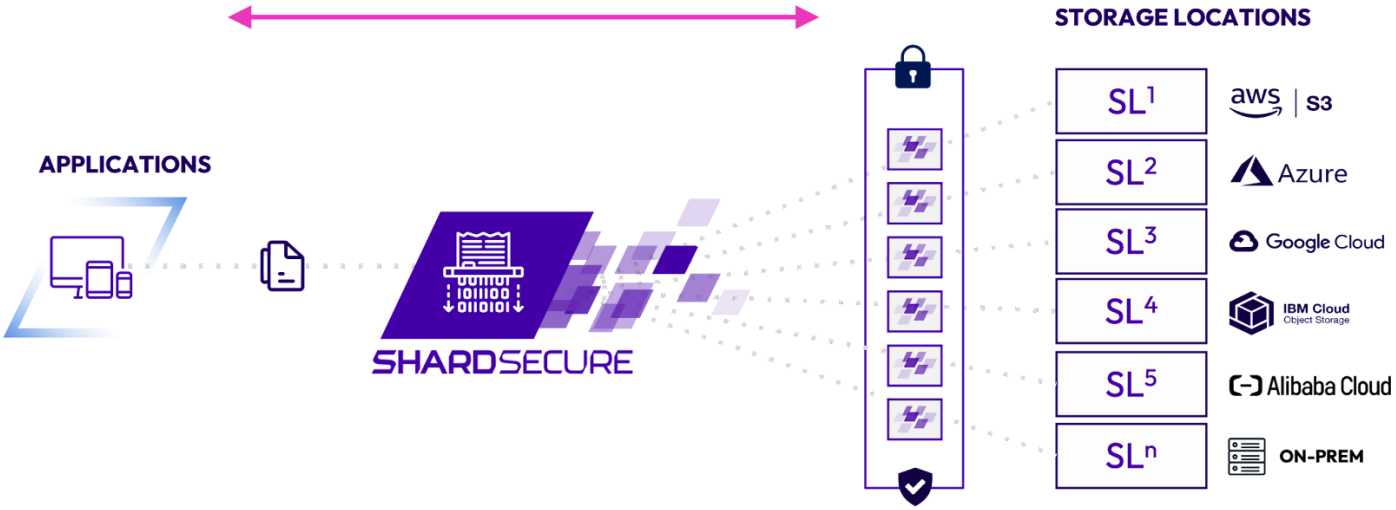
The technology compresses and then digitally shreds data into small fragments known as microshards. The size of the microshards is user-configurable and may be as small as four bytes. The result is that the microshards are too small to contain any personal identifiable information (PII) or other sensitive data.

Mix

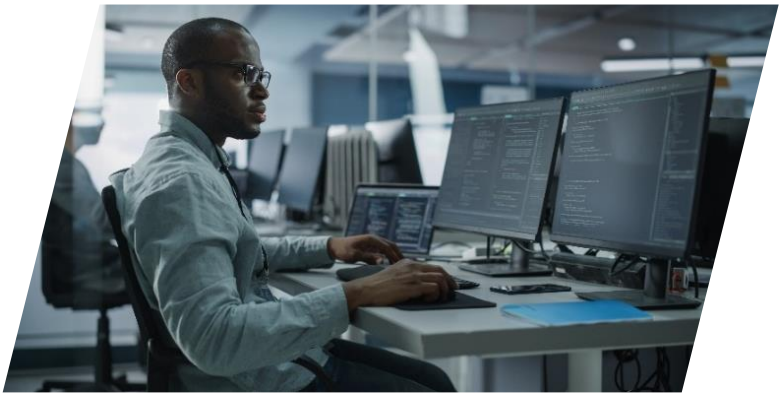
The microshards are next mixed into multiple logical containers known as Microshard containers, which are files containing a subset of microshards. A user-configurable amount of poison data is included to complicate unauthorized reassembly. Additionally, file names, file extensions, and metadata are removed to make unauthorized reassembly virtually impossible.

Distribute

The containers are distributed to multiple customer-owned storage locations. These may be spread across multiple cloud providers, multiple regions under a single cloud provider, and/or hybrid environments.



A Microsharding Example



Consider a 1 MB text file that will be microsharded into four-byte microshards. Depending on the encoding standard used for the file, the result will be over 262,000 one- to four-character microshards — and including poison data will only increase this number. The microsharded data will be mixed into multiple logical containers and distributed across multiple locations (our recommendation would be at least four). No sequential pieces will be stored in the same location.

If an unauthorized user were able to collect all of the microsharded data containers from their various locations for this one file, there would be trillions of possible combinations to try, making it virtually impossible to reconstruct.

Reassembly of Microshard Data

Reassembly of microsharded data is essentially a reversal of the microsharding process. The solution keeps track of all storage locations as well as the order of the microshards for all microsharded data. It is worth noting that the instructions used for reassembly are stored safely in multiple locations with additional security layers, and all the instructions must be used in combination to reassemble the microsharded data. (Please see the “Integrity of the Microsharding Process” section below for more information.)

When a user opens a file, the microsharding engine retrieves the Microshard containers for that file from storage, reassembles the microshards into their proper order while ignoring any poison data, and returns the complete file to the requesting application for the user. This is done in real-time and in parallel and imposes no notable latency.

It is worth noting that each time a file is saved, and therefore microsharded, the order of the microshards and their distribution will change.

Integrity and Availability of Microshard Data

The solution performs multiple data integrity checks and has the ability to reconstruct microsharded data in real-time if any of the microsharded data should fail one of those checks. This capability helps to ensure that any tampering with the data at rest, including encryption by ransomware, is reversed and that business operations can continue unimpacted.

Microsharded data is also able to withstand storage service outages and unauthorized deletion. In both cases, the microsharding engine is able to reconstruct the missing data in real-time, supporting business continuity.

Integrity of the Microsharding Process

A common question is whether an unauthorized third party could install their own instance of ShardSecure in order to reassemble the microsharded data of a targeted organization. The answer is no.

There are three components required to reassemble microsharded data: the locations where the Microshard containers are stored, the correct order of the microshards for any given data, and access to all of the data storage locations. The components that track storage locations and the order of the microshards are encrypted and have additional security safeguards.

Together, these components create a combination that is unique to each installation. Therefore, it is not possible to deploy a new instance of ShardSecure to reassemble the data microsharded by another instance. For instance, if an unauthorized user had gained possession of all Microshard data from customer A, that unauthorized user could not deploy a new instance of ShardSecure and reassemble customer A's microsharded data.

Addressing Use Case 5 with Microshard Technology

ShardSecure Microshard technology is a split processing technology that easily can be deployed in a multi-party processing environment. By using Microshard technology, any organization can process and store any type of data anywhere while staying compliant with the GDPR and the Schrems II ruling.

Below we show how microsharding could be implemented to satisfy the requirements of Use Case 5.

> "The data exporter wishes personal data to be processed jointly by two or more independent processors located in different jurisdictions without disclosing the content of the data to them. Prior to transmission, it splits the data in such a way that no part an individual processor receives suffices to reconstruct the personal data in whole or in part. The data exporter receives the result of the processing from each of the processors independently, and merges the pieces received to arrive at the final result which may constitute personal or aggregated data."

Use Case 5 Requirements:

1

First, you will meet the requirements if a data exporter processes personal data in such a manner that it is split into two or more parts, each of which can no longer be interpreted or attributed to a specific data subject without the use of additional information.

How ShardSecure Helps

Microsharding was created for the express purpose of preventing any unauthorized user or entity from reconstructing or identifying the original data from microsharded data. The function of microsharding is to break data into very small portions (microshards) to rearrange and mix those microshards across multiple logical containers, and to store those containers in multiple customer-owned locations.

The number and location of storage locations is user-configurable; we recommend a minimum of four storage locations. Theoretically, there is no upper limit to the number of storage locations that can be used, though we do recommend a maximum of ten.

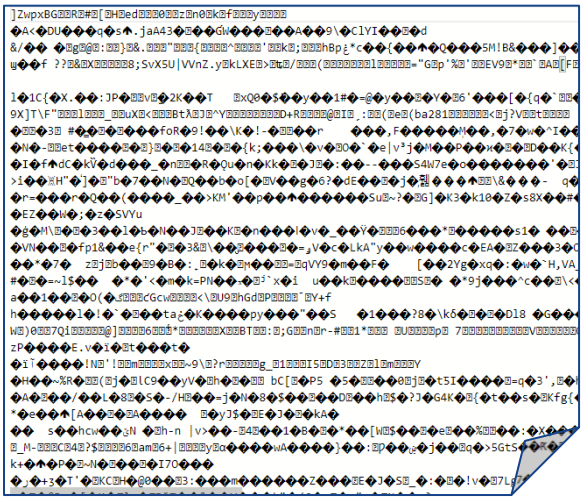


Figure 1: Sample of a container of microsharded data

In the unlikely scenario that an unauthorized user is able to gain access to the microsharded data from every storage location for a given data set, that data set alone is insufficient to reconstruct the microsharded data. There are multiple reasons for this:

1. The microsharding process strips filenames, file extensions, and all other data or metadata that indicates how to reconstruct microsharded data. Each container of microshards is also given a random alpha-numeric name. Therefore, it is not possible to identify which containers were derived from which original data.

2. The original data may be shredded into microshards as small as four bytes, each of which would contain only one to four characters. The size of the microshards is also user-configurable, so it is not possible for an unauthorized user to know how many characters represent a portion of original data.

3. The microsharding process may also use poison or false data to further complicate unauthorized reassembly. Whether to use poison data at all and how much to use is also user configurable.

4. The ShardSecure solution includes a user-configurable policy engine that may be used to customize the size of the microshards, the amount of poison data (if any), the number of containers to use, and more based upon file type and other parameters.

There are multiple components that are kept within the ShardSecure solution that must be used in concert with both each other and the complete data set in order to reassemble any microsharded data. These other components are not stored with the microsharded data. (Please see "Integrity of the Microsharding Process.")

2

You will meet the requirements if each of the pieces is transferred to a separate processor located in a different jurisdiction.

How ShardSecure Helps

The ShardSecure engine policies may be configured to store the Microshard containers in different jurisdictions. These containers may be distributed across different regions of a single cloud provider, across multiple cloud providers, or across a hybrid mix of on-premises storage and one or more cloud providers. This helps to ensure that no one location contains all of the pieces of any given container.

3

You will also meet the requirements if the processors optionally process the data jointly, e.g. using secure multi-party computation, in a way that no information is revealed to any of them that they do not possess prior to the computation.

How ShardSecure Helps

Microsharding was developed to protect data at rest and to keep control of the data in the hands of the data owner. The process, as described throughout this document, ensures that data access by unauthorized users is unintelligible and of no value to those users. Therefore, in the context of our response where the processors are cloud providers, processors will only be granted access to data that the data owner specifically allows through the means that they see fit.

4

You will meet the requirements if the algorithm used for the shared computation is secure against active adversaries.

How ShardSecure Helps

Microshard technology does not use any mathematical computation and therefore cannot be reversed or decoded. Our source code is also protected by microsharding to help prevent supply chain attacks.

5

Additionally, you will meet the requirements if the controller has established by means of a thorough analysis of the data in question, taking into account the missing pieces of information that the public authorities of the recipient countries may be expected to possess and use, that the pieces of personal data it transmits to the processors cannot be attributed to an identified or identifiable natural person even if cross-referenced with such information.

How ShardSecure Helps

The distributed containers have no notion of where the other containers are located, nor do they contain any information on the original data object.

6

Finally, you will meet the requirements if there is no evidence of collaboration between the public authorities located in the respective jurisdictions where each of the processors are located, which would allow them access to all sets of personal data held by the processors and enable them to reconstitute and exploit the content of the personal data in a clear form in circumstances where such exploitation would not respect the essence of the fundamental rights and freedoms of the data subjects. Similarly, public authorities of either country should not have the authority to access personal data held by processors in all jurisdictions concerned.

How ShardSecure Helps

The number and geographic location of the storage locations to which microsharded data will be distributed are both user-configurable and, therefore, within the data owner's control to distribute as they see fit or as is required. It is also possible for data owners to easily move their microsharded data from one storage location to another, whether that be between regions of a single cloud provider, between different cloud providers, or from a cloud provider to on-premises storage.

These capabilities give the data owners the ability to distribute their microsharded data across whichever jurisdictions they see fit.

What is important to note is that, as described throughout this document, microsharded data is unintelligible and of no value to unauthorized users. Merely accessing all of the microsharded data for a given organization is not sufficient to reconstruct the original data. Nor is it possible for an unauthorized user to deploy their own instance of ShardSecure to reassemble a third party's microsharded data. (Please see "Integrity of the Microsharding Process.")

Hans-Peter Erlingsson is the CEO and founder of Lex Legem Advisory & Consulting. Hans-Peter holds a Master of Laws degrees from University of Uppsala and University College Dublin and certified in 2020 as GDPR Privacy Professional. Hans-Peter is specialized in IT law including privacy and data protection law, having 20+ years of experience of privacy and security risk management in different industry sectors. As an example, Hans-Peter played a key role in setting-up and managing the global data privacy compliance program of H&M Hennes & Mauritz alongside with other short- and long-term assignments.

Jesper Tohmo, CTO and Co-Founder. A long-time security expert, Jesper has over 15 years of experience in cybersecurity and cloud computing. Previously, Jesper co-founded security company 2 face commit, served as Cloud Automation Engineer at Scania Sverige and Director of Business Development at McAfee. Jesper also co-founded NordicEdge, a leading identity management company that was acquired by Intel McAfee.

Notice:

This paper is for informational purposes only and creates no legal advice, commitments, nor assurances from ShardSecure, its affiliates, or licensors regarding its products or services. Customers are fully responsible for making their own independent assessments of the information presented herein and for its actions.

Reference documents:

ShardSecure Whitepaper:

https://shardsecure.com/hubfs/resources/white-papers/ShardSecure__White-Paper--Microsharding-Mar-2021.pdf

ShardSecure Microsharding Patent Asset:

https://patents.justia.com/patent/20200143074

ShardSecure Web site:

https://shardsecure.com/

EU GDPR:

https://edpb.europa.eu/system/files/2021-06/edpb_recommendations_202001vo.2.0_supplementarymeasurestransferstools_en.pdf