Secure Stream Processing for Medical Data

info@csem.ch www.csem.ch

Carlos Segarra[1], Enric Muntané[1], Mathieu Lemay[1], Valerio Schiavoni[2], and Ricard Delgado-Gonzalo[1]

[1] *Swiss Center for Electronics and Microtechnology, CSEM SA, Switzerland*
[2] *Université de Neuchâtel, Switzerland*

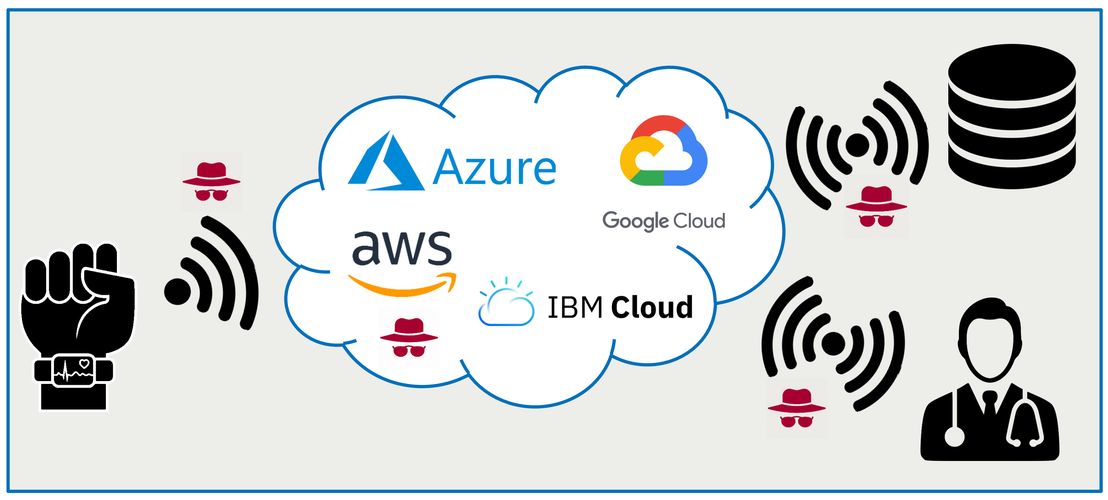
Personalized health and medicine provides the opportunity to benefit from more targeted and effective diagnoses and treatments. To implement it, larger amounts of data and complex processing pipelines are to be implemented. However, recent **data protection regulations** (*e.g.* GDPR) hinder the possibility to leverage powerful computing facilities (clouds) due to **privacy concerns on the user-generated data**.

This poster presents a proof of concept of a streaming IoT architecture that **securely processes cardiac data in the cloud** combining trusted hardware, via **Intel SGX**, and the **Apache Spark** stream processing engine.

Do you trust your cloud provider?

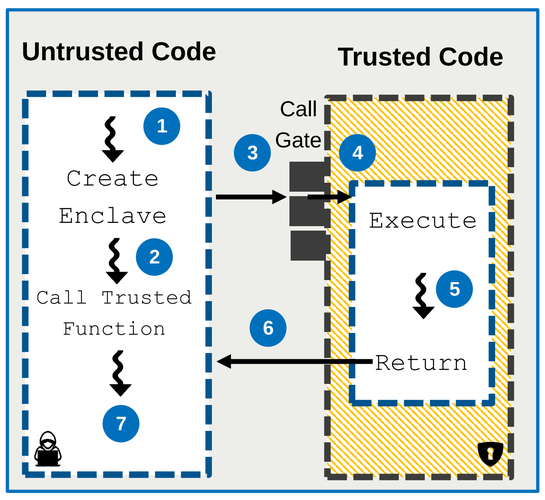
In medical applications…

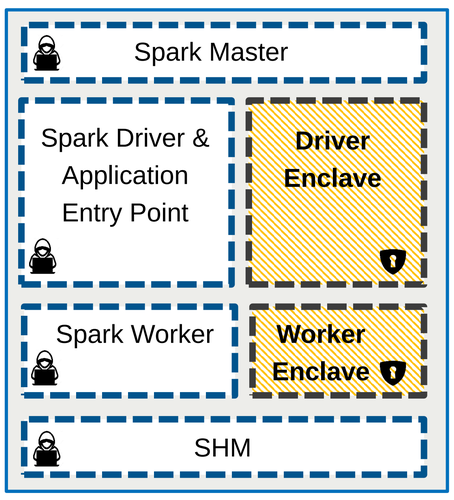
THIRD PARTY CLOUD PROVIDERS SHOULD NOT BE TRUSTED



TEEs, SGX, and SGX-Spark

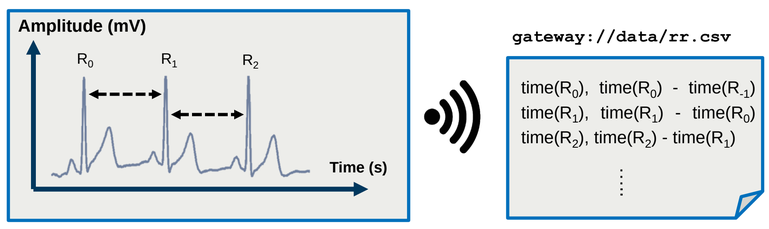
- A **TEE** is a secure area in a processor that grants code and data with **confidentiality** and **integrity.** *E.g.* Intel SGX and Arm TrustZone.
- **Intel SGX** are a set of instructions that enable to create **hardware-protected** areas called **enclaves.**
- **SGX-Spark** is a framework for **seamless** deployment of **Spark** jobs inside **enclaves**.



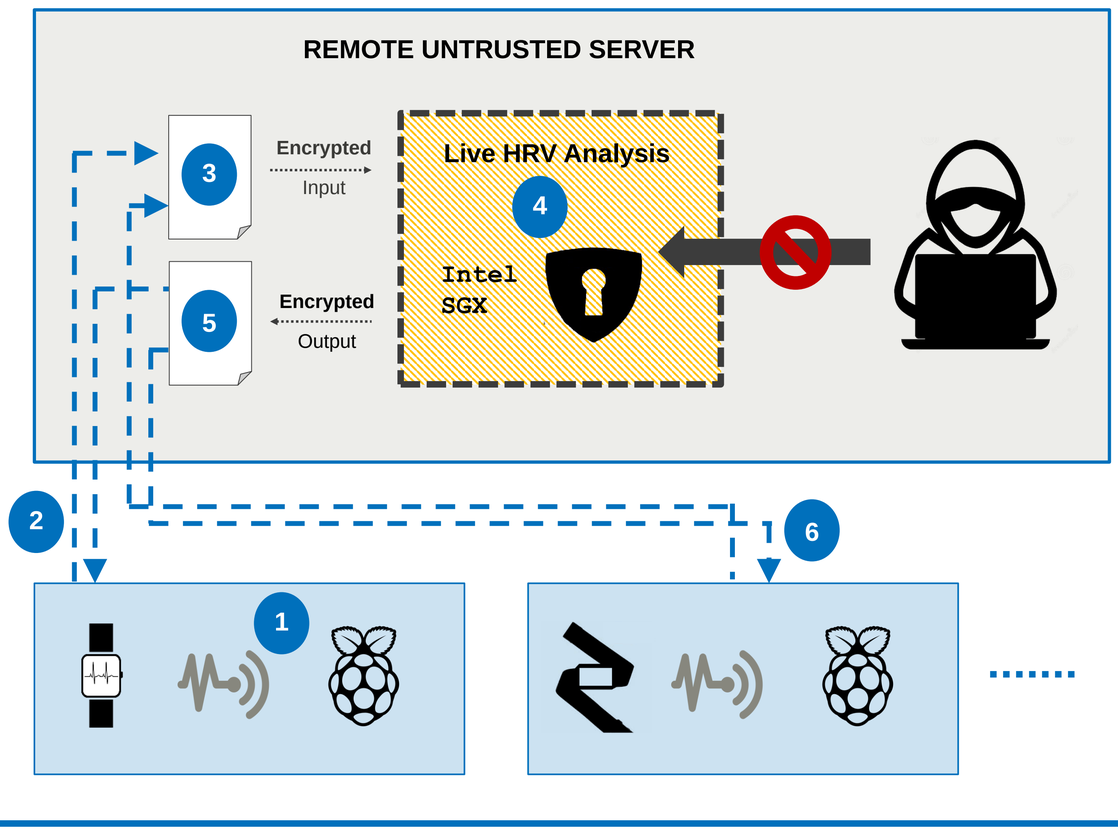


Privacy-Preserving Real Time Cardiac Data Analysis

- We perform real time **Heart Rate Variability (HRV)** analysis.
- We implement:
  - **SDNN**
  - **HRV Band**
- System's **workflow:**



1. ECG/PPG generated at the sensor, **inter-beat intervals** are detected and sent over MQTT to the gateway (RB Pi).
2. Samples are aggregated in files and transferred encrypted from the gateway to the **remote untrusted server** over SFTP**.**
3. Files are stored encrypted in the remote server's filesystem.
4. An **SGX-Spark** streaming job performs live HRV analysis inside enclaves.
5. Results are stored **encrypted** in the remote server's filesystem.
6. Results are transferred from the server to the gateway over SFTP**.**



Evaluation and Results

Evaluation:

- We compare **our solution** with vanilla **Spark Streaming**.
- Evaluation is done in terms of **client** and **load** scalability, until a configuration becomes unstable.
- We consider a configuration becomes **unstable** if the **average processing** time exceeds 10s.

Results:

- Our solution performs computations on untrusted clouds **halving** the system performance.
- It provides **privacy-preserving** data processing **transparently** to the user and developer.