Segurança e Aplicações de Hardware Confiável

Secure Data Lakes

Maria Carreira up202408787
Matilde Simões up202108782
Ricardo Amorim up202107843

Março 2025

Contents

1 Descrição do Problema

2 Visão Geral da Abordagem e do Hardware Confiável Adotado

A utilização de *hardware* confiável é essencial para garantir *data lakes* seguros, especialmente quando se trata de dados sensíveis, como informações médicas. É fundamental garantir a segurança nestes ambientes, onde múltiplos utilizadores precisam do acesso aos dados agregados sem comprometer a privacidade individual. Uma das soluções mais eficazes para este tipo de problema é a utilização do *Intel Software Guard Extensions (SGX)*, que é um ambiente de execução confiável (TEE) que permite o processamento seguro dos dados e protege contra diversos ataque, incluindo ataques provenientes de componentes com altos privilégios do sistema [1].

O *Intel SGX* é uma tecnologia baseada em *hardware* que possibilita a execução de código e o processamento de dados dentro de enclaves protegidos, o que impede acessos não autorizados e modificações indevidas. A escolha do SGX justifica-se pelas seguintes vantagens:

- Confidencialidade: Os dados dentro dos enclaves *SGX* permanecem encriptados e inacessíveis, mesmo para administradores do sistema, prestadores de serviço *cloud* ou atacantes com acesso privilegiado.
- Integridade: O ambiente de execução garante que apenas código confiável é executado dentro dos enclaves, prevenindo alterações não autorizadas.
- Verificação Remota: O *SGX* fornece mecanismos para verificar a integridade e a autenticidade do código em execução dentro de um enclave. Este processo permite que um utilizador verifique a integridade do enclave antes de interagir com ele, assegurando que os dados processados não foram adulterados.
- Minimização da Superfície de Ataque: Ao isolar cálculos sensíveis do sistema operativo, o SGX reduz significativamente o risco de ataques, como a manipulação de memória, injeção de código e explorações ao nível do kernel.

Assim, o *SGX* assegura que os enclaves são isolados e protegidos, permitindo que dados sensíveis sejam processados com segurança dentro do *CPU*, sem ficarem expostos ao software privilegiado [1].

Os dados processados dentro dos enclaves *SGX* estão protegidos contra acessos e modificações não autorizadas devido à estrutura da arquitetura que isola os enclaves em zonas protegidas da memória. Apenas o código autorizado dentro do enclave pode aceder e operar sobre os dados, impedindo que mesmo utilizadores com privilégios administrativos ou atacantes que comprometam o sistema operativo tenham acesso a informações confidenciais. Além disso, as chaves criptográficas utilizadas para proteger os dados nunca saem do enclave, eliminando riscos associados a ataques externos.

A confidencialidade dos dados é assegurada pelo *Memory Encryption Engine (MEE)*, que protege as informações guardadas na *Processor Reserved Memory (PRM)* contra acessos não autorizados [1]. Os dados que se encontram fora do *CPU* permanecem encriptados, garantindo que qualquer tentativa de leitura por parte do sistema operativo

ou atacantes resulte apenas em informação ilegível. Apenas o *CPU*, através do *MEE*, é capaz de realizar a desencriptação em tempo real dos dados, garantindo que o único local onde podem ser lidos em texto claro é dentro do enclave, enquanto estão a ser processados.

A integridade dos enclaves é assegurada pelo mecanismo de verificação remota. Este mecanismo evita ataques como os *rollback attacks*, que tentam reverter um enclave para um estado anterior e potencialmente comprometido que pode conter vulnerabilidades já corrigidas em versões mais recentes. Além disso, a criptografia aplicada pelo *MEE* protege os dados contra ataques de extração de dados sensíveis diretamente da memória RAM, como os *cold boot attacks*. Os dados guardados na *DRAM* não desaparecem imediatamente após se desligar o sistema, logo um atacante pode tentar recuperá-los. Assim, o *MEE* impede que informações sensíveis sejam recuperadas mesmo que um atacante tenha acesso físico ao hardware.

Apesar da segurança avançada oferecida pelo *SGX*, existem potenciais vulnerabilidades que podem ser exploradas para comprometer a sua proteção. Um dos principais desafios são os ataques de *side channels* [2], que podem extrair informações sensíveis ao analisar padrões de acesso à memória, medições de tempo de execução ou consumo de energia. Além disso, falhas no próprio código do enclave podem ser exploradas para executar instruções maliciosas e comprometer a segurança do sistema [3]. Outras ameaças incluem enclaves maliciosos [4], que podem enganar aplicações legítimas e obter dados confidenciais, e ainda falhas na implementação do *SGX*, como demonstrado em ataques como *Foreshadow* [5] `https://www.intel.com/content/www/us/en/security-center/advisory/intel-sa-00161.html`, que explora vulnerabilidades de microarquitetura para aceder a informações protegidas.

No contexto de *data lakes* seguros, o *SGX* possibilita a criação de uma base de dados remota mínima, onde múltiplos utilizadores podem armazenar e consultar informações com garantias de privacidade. Esta abordagem permite que consultas, como a análise de dados médicos, sejam processadas dentro dos enclaves, assegurando que apenas resultados agregados, como médias estatísticas, sejam expostos, sem revelar dados individuais. Além disso, facilita a partilha e colaboração segura entre diferentes entidades, como hospitais e instituições. Por fim, ao ser implementado em ambientes de *cloud*, o *SGX* protege as informações contra ameaças internas, garantindo que nem os próprios fornecedores de serviços conseguem aceder aos dados guardados.

Concluindo, a utilização do *Intel SGX* neste projeto oferece garantias sólidas de segurança, confidencialidade e integridade. Através da aplicabilidade em ambientes de execução confiáveis, conseguimos implementar o armazenamento e o processamento de dados sensíveis e ainda permitir a partilha de dados sem comprometer a privacidade. Esta abordagem garante que os dados permanecem sempre protegidos, mesmo em ambientes sujeitos a possíveis ameaças. Com os mecanismos de enclaves e verificação remota, o *SGX* consegue garantir robustez para proteger *data lakes*, especialmente em áreas críticas como a saúde.

3 Abordagem Proposta

3.1 Visão Geral da Solução

3.2 Entidades do Sistema

3.3 Arquitetura Geral

4 Requisitos Funcionais e de Segurança do Projeto

5 Descrição do progresso feito e Próximos objetivos

References

[1] Daniel Ehnes. The magic of intel's sgx. a tutorial on programming a secure enclave, 2018.

[2] Ferdinand Brasser, Urs Müller, Alexandra Dmitrienko, Kari Kostiainen, Srdjan Capkun, and Ahmad Reza Sadeghi. Software grand exposure: Sgx cache attacks are practical. *11th USENIX Workshop on Offensive Technologies, WOOT 2017, co-located with USENIX Security 2017*, 2 2017.

[3] Michael Schwarz, Samuel Weiser, and Daniel Gruss. Practical enclave malware with intel sgx. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11543 LNCS:177–196, 2 2019.

[4] JP Aumasson and Luis Merino Kudelski Security. Sgx secure enclaves in practice: Security and crypto review. 2016.

[5] Jo Van Bulck, Marina Minkin, Ofir Weisse, Daniel Genkin, Baris Kasikci, Frank Piessens, Mark Silberstein, Thomas F Wenisch, Yuval Yarom, Raoul Strackx, and Ku Leuven. Foreshadow: Extracting the keys to the intel sgx kingdom with transient out-of-order execution.