IronSafe
A Secure and Policy-Compliant Query Processing Architecture

Harshavardhan Unnibhavi

David Cerdeira, Antonio Barbalace, Nuno Santos, Pramod Bhatotia

ACM SIGMOD 2022
Data analytics in the cloud

Cloud enables scalable and fault-tolerant computing in a cost-effective manner
The challenges of cloud computing

#1: Security
#2: Policy compliance
How do we ensure security in untrusted cloud environments?
How do we ensure policy-compliant query processing that is auditable by a regulator?
Problem statement

To design a **secure, policy-compliant and high-performant** query processing architecture for untrusted cloud environments.
Our proposal

IronSafe
A secure and policy-compliant query processing architecture

Security
Policy compliance
High performance
Key building blocks

IronSafe

Hardware-assisted trusted computing + Policy language and compliance infrastructure + Near data processing (NDP)
Outline

- Motivation
- Background
- Design challenges
- Workflow
- Evaluation
Trusted computing

● Abstraction
  ○ Trusted Execution Environment (TEE)
  ○ Rich Execution Environment (REE)

● Remote attestation
  ○ Authenticate hardware and software

IronSafe leverages trusted computing to provide strong security guarantees
Near data processing (NDP)

- Heterogeneous deployment
  - Specialized accelerators for data processing
  - x86 host and ARM storage system

- Near data processing
  - Offload computation to storage

IronSafe leverages NDP to achieve high performance
Outline

• Motivation
• Background
  • Design challenges
  • Workflow
  • Evaluation
Challenge #1: Heterogenous TEEs

- Heterogeneous host and storage
  - Heterogeneous ISAs and TEEs
  - Different security guarantees

- Combine execution across TEEs
  - Host: Intel SGX
  - Storage: ARM TrustZone

A heterogeneous confidential computing framework
#1: A heterogeneous confidential computing framework
Challenge #2: Policy specification and compliance

- **Policy specification**
  - Expressible but simple and less error-prone

- **Policy enforcement**
  - Guarantee integrity of all components for every query
  - Multiple versions of hardware and software

- **Proof of policy compliance**
  - Guarantee compliance for every query

A policy compliance monitor
#2: Policy compliance monitor

- **Policy specification**
  - IronSafe’s declarative policy language

- **Policy enforcement**
  - Attestation of both host and storage system

- **Proofs of policy compliance**
  - Cryptographic certificates signed by trusted monitor
Outline

• Motivation
• Background
• Design challenges
  • Workflow
  • Evaluation
IronSafe workflow

#1
Trust establishment

#2
Policy compliance

#3
Query execution
#1: Trust establishment

![Diagram](image)

- **Client**
  - Query + Policy

- Host
  - Enclave
    - Query engine

- Storage system
  - REE
    - Storage-side query engine
  - TEE
    - Secure storage
    - Attestation

- Trusted monitor
  - Trust establishment

- Attestation
#2: Policy enforcement

- **Client**
- **Query** + **Policy**

**Host**
- Enclave
- Query engine

**Storage system**
- **REE**
  - Storage-side query engine
- **TEE**
  - Secure storage
  - Attestation

**Storage medium**

**Trusted monitor**
- Policy enforcement
- Trust establishment

**Certificate**
#3: Query execution

- Client
  - Query
  - Policy
  - Results
  - Compliance certificate

- Host
  - Enclave
    - Query engine
    - Query offload
    - Results

- Storage system
  - REE
    - Offloaded query
    - Storage-side query engine
  - TEE
    - Secure storage
    - Attestation
  - Storage medium

- Trusted monitor
  - Policy enforcement
  - Trust establishment
Outline

- Motivation
- Background
- Design challenges
- Workflow
- Evaluation
Evaluation

- Questions
  1) Performance gains
  2) Policy compliance use cases

- Workloads
  - SQL queries: TPC-H benchmark
  - GDPR anti-patterns (policy compliance)

- Setup
  - Host: i9-10900K, 64GiB memory
  - Interconnect: 40GbE
  - Storage: 16 core ARM cortex A72, 32 GiB memory, 1TB M.2 NVMe
Q1: Performance gains

IronSafe speedup compared to secure host only processing

IronSafe is 2.3x faster on average and guarantees strong security and policy compliance
Q2: Policy compliance use cases

<table>
<thead>
<tr>
<th>GDPR Anti-patterns[^1]</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timely deletion</td>
<td>Ensure data is not stored and used indefinitely</td>
</tr>
<tr>
<td>Indiscrimination</td>
<td>Control data sharing</td>
</tr>
<tr>
<td>Transparency</td>
<td>Transparent sharing of data</td>
</tr>
<tr>
<td>Risk agnostic</td>
<td>Prevent untrusted parties from accessing data</td>
</tr>
<tr>
<td>Data breaches</td>
<td>Know if data was used by unwanted parties</td>
</tr>
</tbody>
</table>

Summary

• Cloud systems enable high-performance query processing

• Processing queries in a secure and policy compliant manner is challenging
  • Heterogenous TEEs + Policy specification and compliance

• Ironsafe\textsuperscript{[1]}: A secure and policy compliant query processing architecture
  • A heterogenous confidential computing framework
  • IronSafe’s declarative policy language
  • A compliance infrastructure for enforcing policies

https://github.com/harshanavkis/ironsafe
Backup slides
Threat model

- Host
  - Host memory
  - Query engine
- Interconnect
- Storage system
  - Offloaded query
  - Storage medium
IronSafe overheads

Overheads due to providing confidentiality, integrity and freshness guarantees
Timely deletion of data

read :- sessionKeyls(K_A) | sessionKeyls(K_B) & le(T, TIMESTAMP)