Privacy-preserving collaboration

Basma El Gaabouri, Christopher Haster, Derek Miller, Dominic Mulligan, Nick Spinale, Shale Xiong

Systems Group, Arm Research
Background

We believe that **strong isolation technology** and **remote attestation**:

- Allow the design of novel data-intensive applications with fine-grained access control,
- Allow computations to be safely moved around, without sacrificing privacy or integrity,
- Potentially separate *possession* of data from *control* over that data

Here, **strong isolation** is our term for a range of hardware- and firmware-based isolation mechanisms, aiming to provide strong privacy and integrity guarantees

*Veracruz* is our vehicle for understanding what these technologies are capable of
The Veracruz framework
A framework for defining flexible and efficient multi-party computations

Veracruz aims to support common use-cases for advanced cryptographic techniques
• Techniques like homomorphic encryption, secure-multiparty computations, and similar

Unlike those techniques, we aim to be:

1. **Efficient**: Be fast enough to execute “interesting” programs,
2. **Familiar**: Allow programmers to use familiar programming languages and tools,
3. **General**: Seamlessly support a large class of multi-party computations,
4. **Reusable**: Provide a single framework supporting a wide-range of privacy-preserving computations without requiring significant reconfiguration for each task

In common with those techniques, we aim to provide a strong **security/privacy guarantee**
Veracruz from 50,000ft

The **data inputs** to Veracruz. Note that these can originate from different agents who are mutually distrusting.
Veracruz from 50,000ft

The program, which may originate from an agent distinct from those providing the data inputs. In Veracruz, we use WebAssembly (WASM) as our executable.
Veracruz from 50,000ft

Program \( P \)

Data_1 \quad Data_2 \quad Data_N

Program and data are provisioned securely into Veracruz, running on a host, which computes a result by applying the program to the data.
A policy details the roles and identities of all involved in the computation and describes who can retrieve the result.
Veracruz from 50,000ft

To maintain secrecy we need to control the **expressivity** of the program $P$, and the **capabilities** of its environment, which computes the result.

$$P(Data_1, Data_2, \ldots, Data_N)$$
Use-case: privacy-preserving machine learning
Use-case: privacy-preserving machine learning
Use-case: privacy-preserving machine learning
Use-case: privacy-preserving machine learning
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

Veracruz

Global policy
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

A45B3201: £4.99
E3332110: £34.23
01224573: £17.50
...
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

<table>
<thead>
<tr>
<th>ID</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>A45B3201</td>
<td>£4.99</td>
</tr>
<tr>
<td>E3332110</td>
<td>£34.23</td>
</tr>
<tr>
<td>01224573</td>
<td>£17.50</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

A45B3201: £4.99
E3332110: £34.23
01224573: £17.50

Veracruz

Global policy
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

A45B3201: £4.99
E3332110: £34.23
01224573: £17.50
...

Veracruz

Global policy
Use-case: privacy-preserving set-sum computation

Internet advertising platform

Client

A45B3201: £4.99
E3332110: £34.23
01224573: £17.50
...

Σ referred customer spend
...and many more potential use-cases

1. IP protection,
2. Privacy-preserving surveys/auctions/elections,
3. Privacy-preserving distributed compute: map-reduce/grid computing *a la* SETI@home,
4. Private search/fuzzy matching,
5. Provenance tracking for data,
6. Verifiable computation,
7. N-way secret sharing,
8. Fair exchange of documents,
9. Zero-knowledge proof of knowledge,
10. Delegating computations from weak devices to untrusted servers,

...*ad infinitum*
Abstracting over isolates

Veracruz supports *multiple* different isolation technologies at present:

- **Arm TrustZone** trusted applications,
- **Intel SGX** secure enclaves,
- The high-assurance **seL4 microkernel**,
- **AWS Nitro Enclaves**, ...and maybe more in the future, representing different points on a *continuum of paranoia*

Veracruz provides abstractions over isolate technologies, with:

- A single, portable programming model based on **WebAssembly**,  
- A unified attestation mechanism, based on **Arm’s PSA Attestation** protocol, which hides platform-specific attestation protocols from clients
A few future directions

• Support for streaming computations

• Adoption of a subset of WASI as our ABI

• Multi-isolate use-cases, e.g. privacy-preserving grid-compute, or map-reduce

• Dynamic checking of the runtime behaviour of the program

• Supporting more isolation technologies
Conclusions

Veracruz is a research project exploring how strong isolation technology and remote attestation can influence the design of novel, data-intensive distributed systems.

Veracruz allows users to easily design and deploy collaborative, privacy-preserving computations using a range of software and hardware isolation mechanisms and WASM:

- Arm TrustZone trusted applications,
- Intel SGX enclaves,
- The seL4 high-assurance hypervisor,
- AWS Nitro Enclaves.

Veracruz has many potential applications, which we are only just beginning to explore!
Get involved

Veracruz is (provisionally) adopted as a project by the Confidential Compute Consortium, and all of our development is now out in the open, on Github:

https://github.com/veracruz-project/veracruz

We are interested in attracting collaborators to help us drive the project forward. If you’re interested in getting involved, e-mail any of the team members or get in touch via Github!