ELASTIC Project
Secure Systems Group, Aalto University

Quang Huy Nguyen, Huy Ngo, Lachlan J. Gunn

Isolating high-performance ML workloads with WebAssembly

WebAssembly Components

Portable and sandboxed workloads with modular interfaces and lightweight isolation.

Confidential Virtual Machines (CVMs)

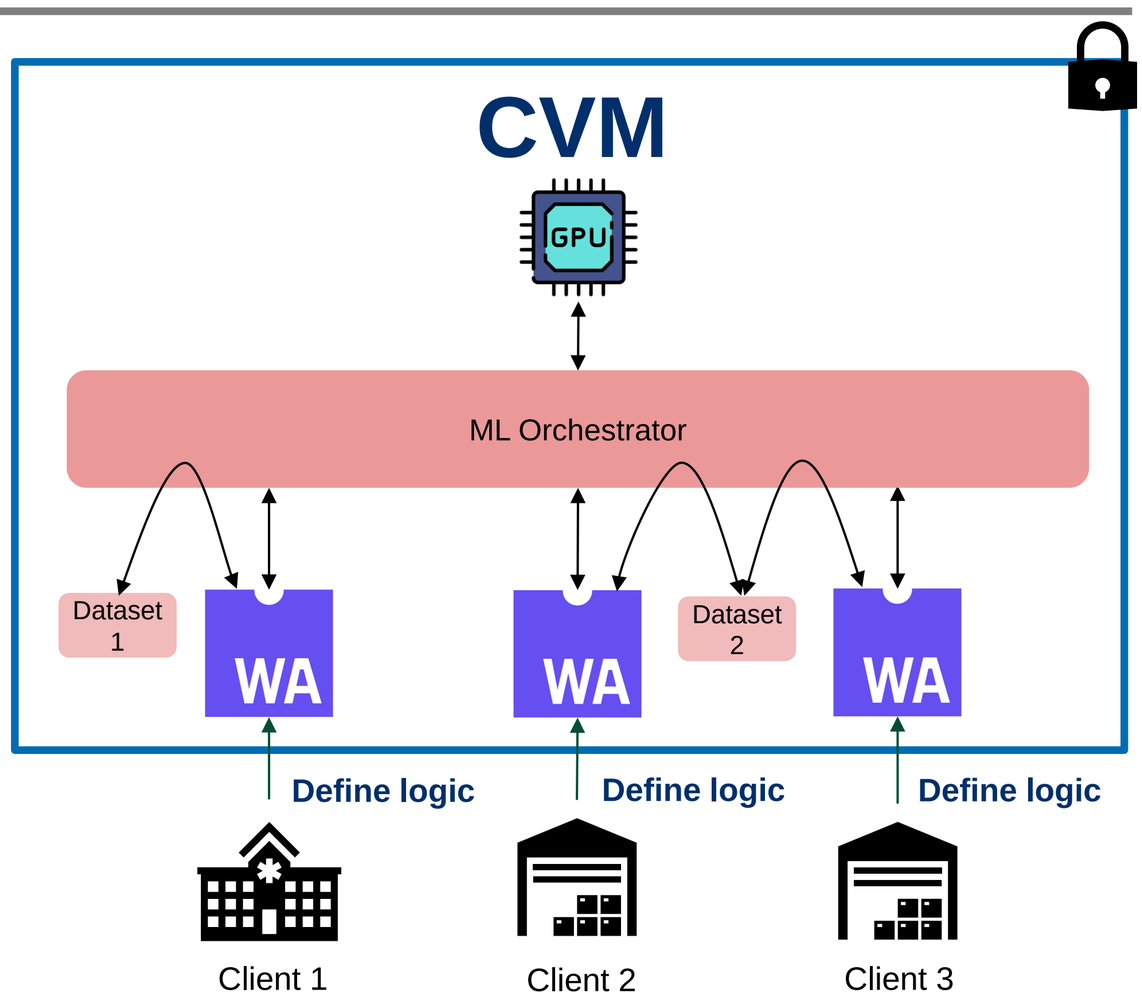
Trusted virtual machines with hardware-enforced memory isolation and remote attestation capabilities.

Motivations

- CVMs protect sensitive computations and data.
- Confidential GPU provide limited partitioning capabilities (e.g., NVIDIA H100 MIG supports up to 7 GPU instances).

Design Goals

- ❖ Low-cost isolation within CVMs.
- ❖ Workload flexibility through Wasm-managed logic.
- ❖ Compatibility with GPU accelerators.
- ❖ Remote attestation should also allow you to verify the results

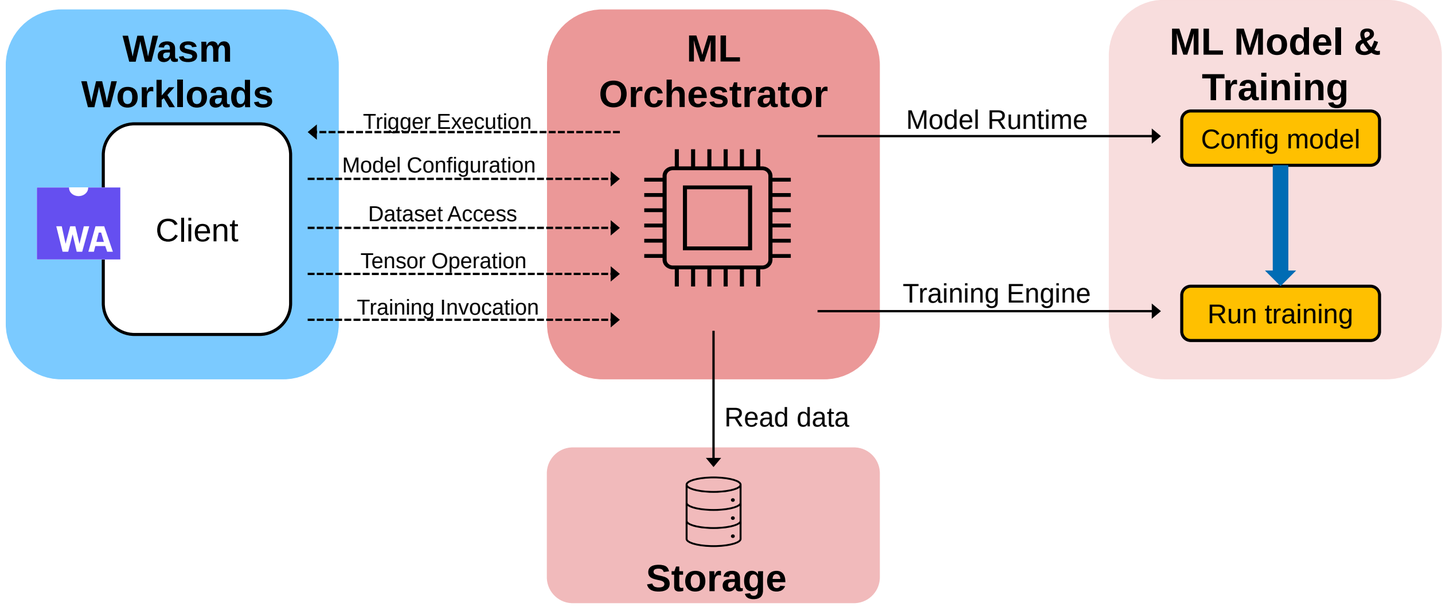


Implementation

- Design a generic WIT-based interface to orchestrate ML workflows across the Wasm boundary :
  - WIT types for models, tensors, and datasets
  - Resource handles to avoid copying data in and out of the sandbox
- Host implementation in Rust, using the `tch` Torch bindings for high-performance ML execution.

Performance

Data size	Type	Training	Pre-computation	Other	Total time
1GB	Native	86.70s ± 0.16s	2.58s ± 0.07s	1.05s ± 0.01s	90.33s ± 0.17s
	Wasm	87.35s ± 0.15s	3.33s ± 0.10s	1.10s ± 0.01s	91.78s ± 0.16s
	Overhead	**0.75% ± 0.26%**	**29.01% ± 1.09%**	**4.70% ± 1.02%**	**1.60% ± 0.26%**
3GB	Native	282.47s ± 0.46s	6.71s ± 0.05s	1.32s ± 0.02s	290.50s ± 0.44s
	Wasm	284.50s ± 0.46s	8.75s ± 0.08s	1.38s ± 0.01s	294.62s ± 0.45s
	Overhead	**0.71% ± 0.23%**	**30.32% ± 1.45%**	**4.55% ± 1.42%**	**1.42% ± 0.22%**



Future work

- Integrate CVM/TEE to enable attestation for verifiable ML.
- Implement different ML workloads (federated learning, generative AI, inference, etc.)

Aalto University

Co-funded by the European Union

The ELASTIC project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe programme, Grant Agreement No. 101139067. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or SNS-JU. Neither the European Union nor the granting authority can be held responsible for them.