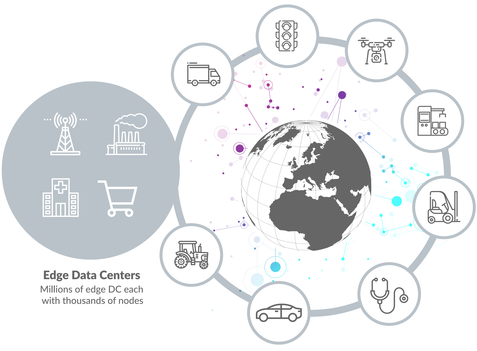
SimpleMachines

AI's Impact on Hardware:

Why Chips Need To Be Redesigned to Keep Up With New Neural Networks

Artificial Intelligence (AI) and Machine Learning (ML) is continually evolving, and chip design must evolve accordingly in order to successfully enable and support The Age of AI.

The Age of AI: A Trillion Connected Things



2M Factories
570M Farms
40M Miles Of Road
13M Stores
5M Call Center Agents
10T Miles Driven Per Year

Houston, we have a problem!

Artificial Intelligence (AI) and Machine Learning (ML) is continually evolving, and chip design must evolve accordingly in order to successfully enable and support The Age of AI.

Present and future machine learning algorithms demand new paradigms in compute capabilities.

Existing CPUs and GPUs are programmable but do not provide the necessary computational efficiency required for AI time-scalability.

Solutions focusing on computational efficiency for AI algorithms lack programmability.

At the rapid pace at which AI algorithms continually evolve, these solutions are virtually obsolete on arrival.

The new AI compute solutions coming to the market offer neither programmability NOR time-scalability.

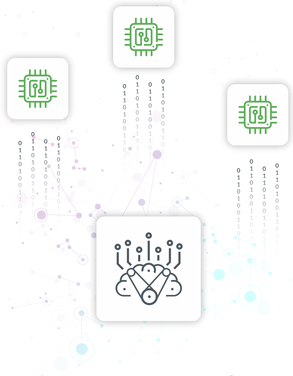
Composable Computing: The paradigm shift

SimpleMachines delivers a powerful breakthrough platform that marries computation efficiency with programmability and time-scalability.

The hardware paradigm shift is in the design of Composable Behavior Execution: a clean-slate design breaking away from 30 years of incremental innovations. Traditional CPUs execute one "instruction" or line-of-code at a time, where the chip has no knowledge of data or the global scope of this instruction's role in the entire program.

SimpleMachines chip instead directly manipulates and understands program properties: data size and shape, whole program size, and shape.

With this global information, our software stack transforms the chip's storage and execution mechanisms on-the-fly to match the applications' data and computation patterns, achieving the same effect of having a custom chip built for that application.



These ideas came out of:

64 person years of research

6 PhDs

7 best-paper awards

13 patents

20 invention disclosures

A clean slate chip design

SimpleMachines' compiler and chip hardware is based on identifying four fundamental behaviors that are universal and central to many algorithms:

1 - Operand communication,
2 - Synchronization,
3 - Computation, and
4 - Control

Our chip implementation directly implements these four behaviors, creating an engine that runs as efficiently as a customized chip.

This allows us to create a platform that is completely under software control, while running at the efficiency of a fully customized chip designed for those applications.

Advantages

Flexible

Flexibly implement, decode, execute, and writeback as coarse-grained blocks on chip through our composable behavior engine architecture. The program's machine code provides a list of functions in the sequence that they happen, be it at the same time or independently one after the other. For any application, the relative balance and interaction between these behavior changes are controlled and orchestrated by our dynamic run-time engine.

Real-time

Leveraging the advances in machine learning (integer linear programming), compiler technology, and chip architecture, automate the design, implementation, and synthesis tasks on-the-fly in real time with our software stack—something that an ASIC designer achieves in months for custom chips.

Powerful

Deliver the power of custom chips for custom applications—without the traditional time-to-market obsolescence, and the expensive production costs associated with ASICs.

SimpleMachines

SimpleMachines.ai

https://www.simplemachines.ai/contact

https://www.linkedin.com/company/simplemachines-inc/

2890 Zanker Rd, Suite 120, San Jose, CA 95134