

TOWARDS ZERO-WASTE COMPUTING

Jeffrey Spaan, Kuan-Hsun Chen, Ana-Lucia Varbanescu
CAES, EEMCS, UTwente

The energy consumption of computing is substantial and constantly increasing!

Stakeholders and call(s) to action(s)



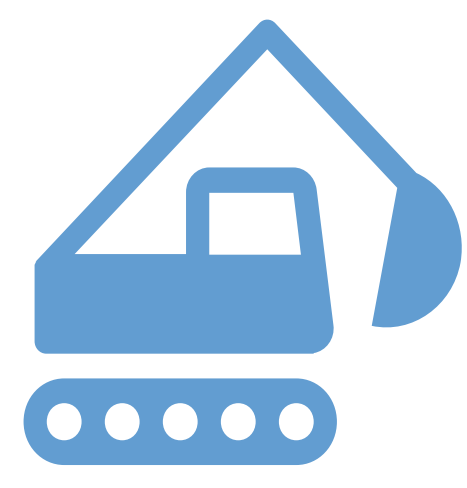
Developers and users

Improve the energy efficiency of their codes, making use of algorithmic, programming, and hardware tools.

Design and implement applications able to adapt to the available system resources



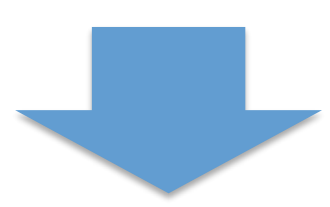
Waste detection & Efficiency improvement



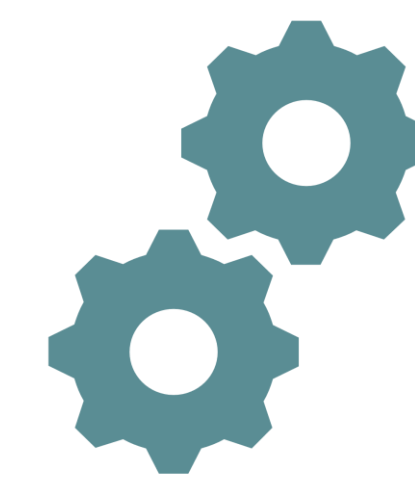
System operators

Ensure efficient scheduling of workloads on system resources.

Harvest energy where resources/systems are massively underutilized.



Efficient operation & novel resource management



System integrators

Offer the right mix of resources for the application developers and system operators.

Include efficient hardware to enable different application mixes.



Sustainable acquisition & extended lifetime

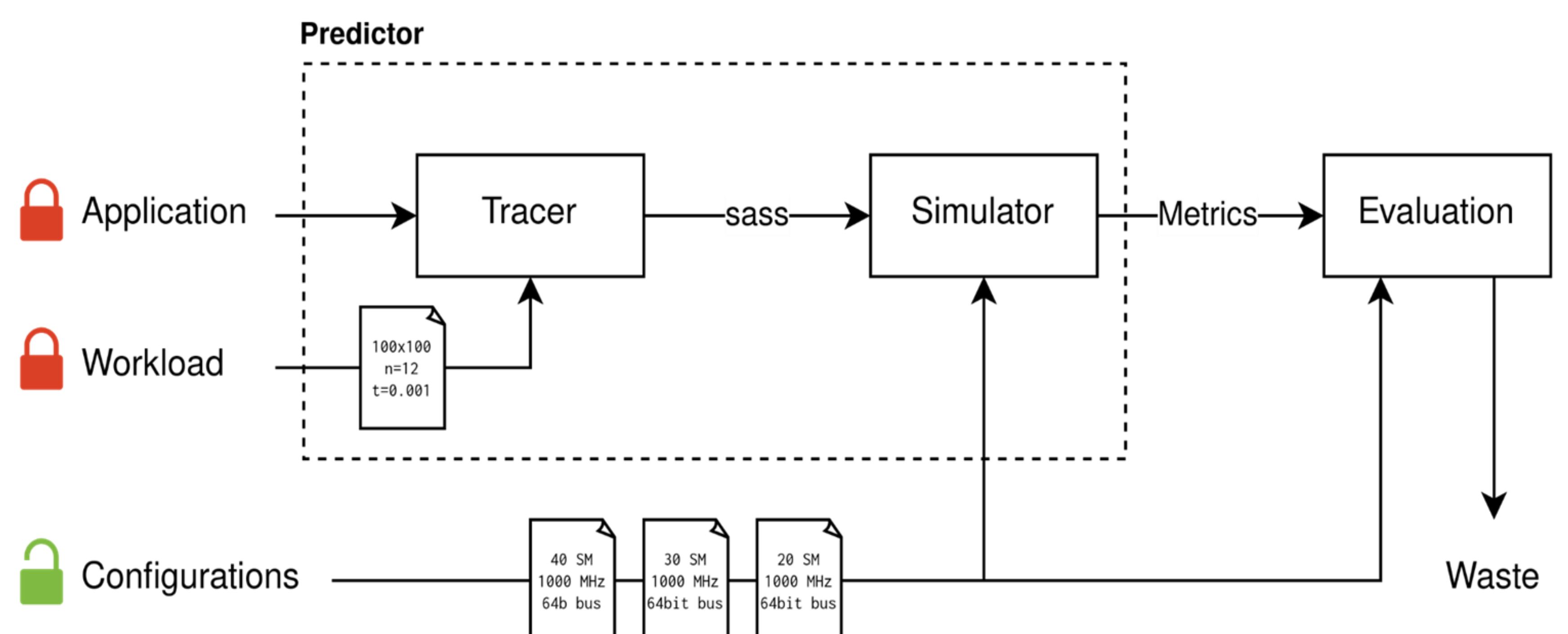
Unnecessary time (or energy) spent in (inefficient) computing is **compute waste**.
Is there waste in computing systems & applications?

Workflow

Approach: To detect waste, we shrink the given hardware platform and observe performance.

Expectation: Constant performance \leftrightarrow waste in the original configuration.

Validation: 5 applications, multiple workloads, one GPU (RTX 2060 Super, AccelSim), various configurations (different SMs, different clock frequencies)



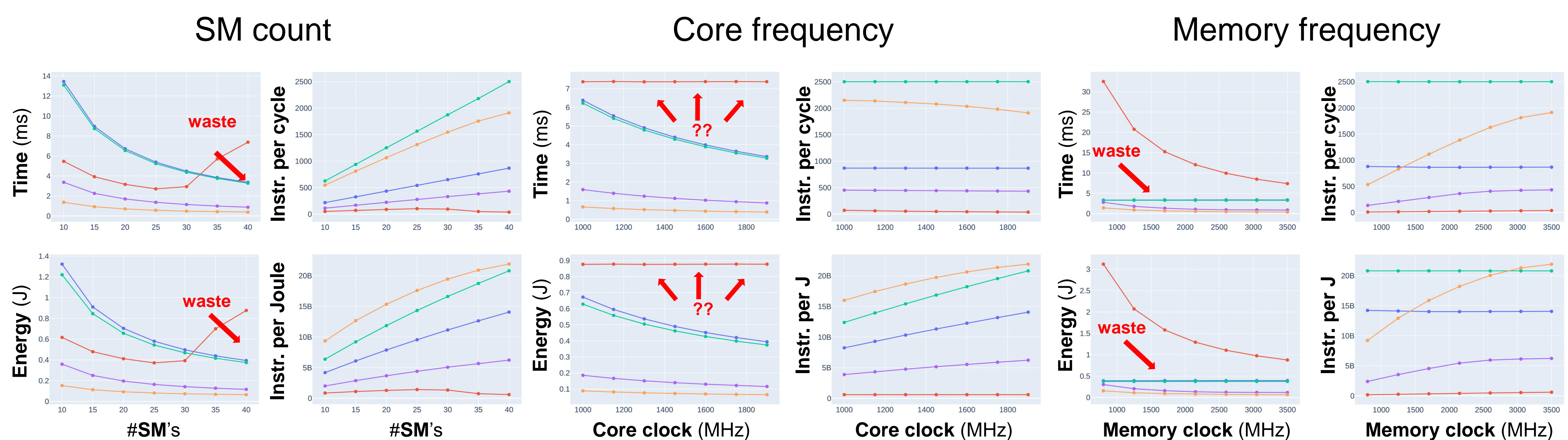
Results

Compute-bound

- Hotspot
- K-means (2)

Memory-bound:

- K-means (1)
- Backprop (1)
- Backprop (2)



There exists waste due to the mismatch between the application requirements and system configuration.

We work on new applications and systems, and better ways to adapt systems to applications requirements.