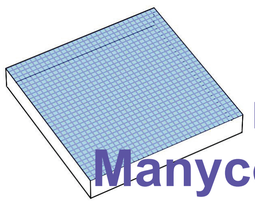


Highly Parallel
Rad-Hard
RISC-V-based
Manycore Accelerators

Luca Benini, Ran Ginosar, Frank K. Gürkaynak, Pablo Ghiglino, David Steenari

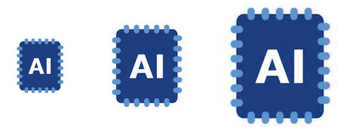
RISC-V in Space workshop, April 2025

MOTIVATION: compute

- Space AI requires ever growing compute & storage
  - Analytics: classification, detection, information extraction, data selection
  - Predictive: what's next
  - Autonomy: decision making
  - RAG: domain-specific reasoning & autonomy
  - Agentic AI: assisted by classical algorithms, esp. DSP

MOTIVATION: flexibility

- Space AI requires flexibility
  - Enable evolution over mission lifetime
    - future types of AI
  - Enable multiple applications & use cases, multiple customers



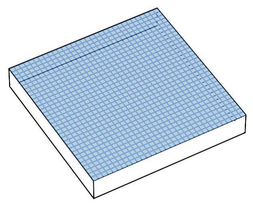
MOTIVATION: acceleration

- Space AI calls for an accelerator
  - Rad-hard multicore CPUs are:
    - not powerful enough,
    - power-hungry,
    - not scalable

  - Rad-hard FPGAs not flexible, hard to program
  - COTS GPUs are not rad-hard, hard to program
  - COTS NPUs are not rad-hard, are inflexible

MOTIVATION: parallel

- Space AI accelerator should be **manycore**
  - Replacing powerful single core CPU by m tiny simple cores on same area:
    - execute $\sqrt{m}$ faster ('speedup')
    - consume $1/\sqrt{m}$ power
    - consume $1/m$ energy

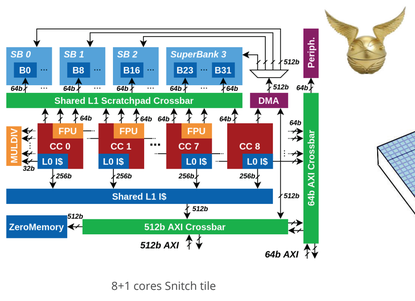


MOTIVATION: resources

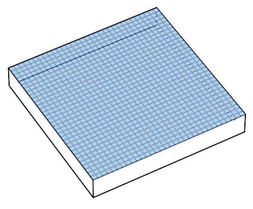
- Space market is small
  - No commercial ROI expected
    - Full ESA funding required
  - No budget for expensive chip design
  - No budget for complex programming
  - One solution for all Space use cases

ARCHITECTURE: many cores

- Manycore ≡ 1024 or more cores on chip
  - small cores, such as RISC-V Snitch

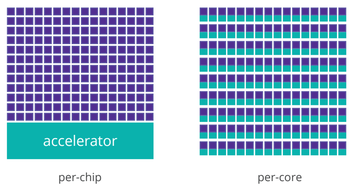


8+1 cores Snitch tile



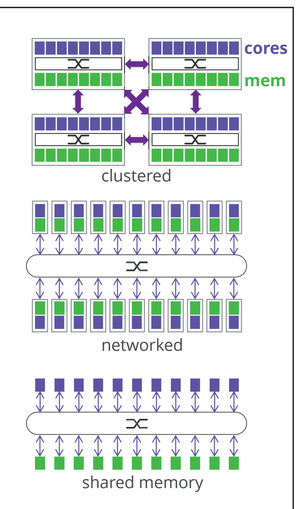
ARCHITECTURE: per-core/per-chip

- Per-core vs. per-chip acceleration
  - Accelerate AI & DSP operators: MM, SoftMax, GELU
  - Trade off:
    - Number of cores
    - Size if cores
    - Size of per-chip accelerator
  - Trade off parameters:
    - area,
    - performance,
    - power,
    - ease of programming



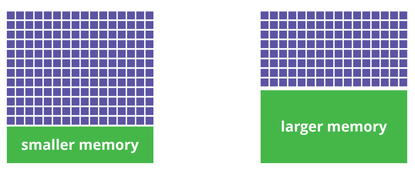
ARCHITECTURE: alternatives

- Alternative architectures, programming models
  - **Clustered**
    - Challenge: access across clusters
  - **Networked**
    - Challenge: message-passing model
  - **Shared memory**
    - Challenge: access of remote data
  - Data caches vs. uncached memory
  - Compiler-generated pre-fetch by non-blocking access
    - Cache or score-boarding



ARCHITECTURE: memory vs. cores

- Size of on-chip memory vs. number of cores
  - Requires simulation studies and optimization
  - Factor: size of caches and scratchpad per core
  - Shared memory as LLC vs. Main & DMA

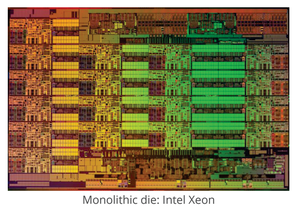


ARCHITECTURE: RISC-V ISA

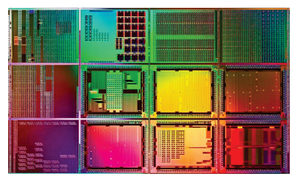
- RISC-V ISA is good for AI & DSP
  - ISA extensions considered:
    - Support linear algebra, varying bit precisions, non-linear operators
    - Support acceleration

ARCHITECTURE: monolithic vs. SiP

- Monolithic die vs. homogeneous SiP with multiple chiplets
  - Monolithic die
    - performs better
    - NRE cost is higher
  - SiP with chiplets
    - NRE is lower, TTM is shorter
    - re-use other chiplets: interface, CPU, FPGA, …
    - incurs D2D delays & power



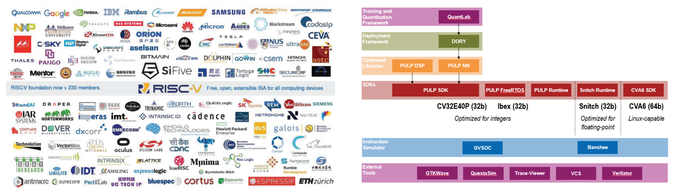
Monolithic die: Intel Xeon



chiplets

SOFTWARE: RISC-V ecosystem

- Rich ecosystem for RISC-V
  - Compilers, profilers, simulators, analyzers, frameworks, runtime managers
  - Parallel / manycore languages (OpenMP, MPI, OpenCL, C+TOP)

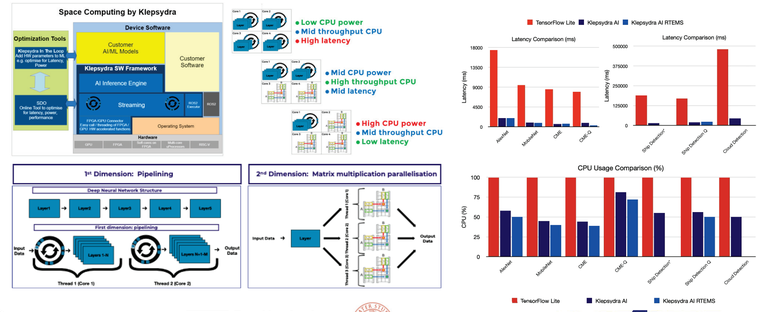


SOFTWARE: evolving AI

- Support evolving AI/ML:
  - Analytics
    - ship / fire detection, cloud-screening, …
  - Predictive analysis
    - fault estimation, debris threat, …
  - Autonomy
    - preventive maintenance, constellation self-management, …
  - Evolve from DL to LLM, transformers, Retrieval-Augmented Generation (RAG), Agentic AI

SOFTWARE: AI Infrastructure

- Klepsydra AI and Streaming framework optimize acceleration on CPUs & manycores

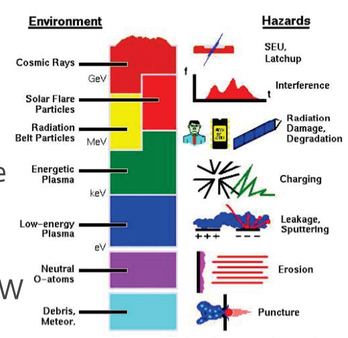


SOFTWARE: certification

- Challenge: ESA Class Alpha missions require rigorous software certification

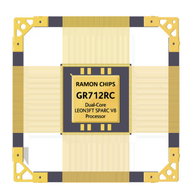
SPACE: challenges

- Space challenges
  - Radiation hardening: TID, SEL, SEU
  - Thermal resilience: thermal cycles, temp range
  - Mechanical resilience: shock, vibrations
  - Mitigation by FDIR HW & SW



SPACE: quality

- Quality range: from short lifetime LEO to ESA Class Alpha
  - Addressed by a variety of packages for different segments



Class S/V package

LEO package