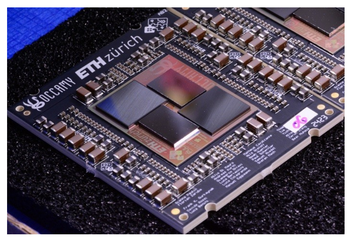
321

HCI: A Flexible and Scalable Interconnect for Tightly-Coupled Heterogeneous Accelerator Clusters

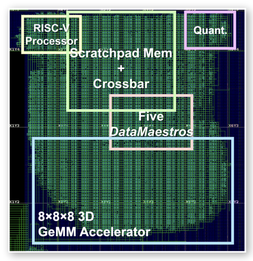
Sergio Mazzola, Arpan Suravi Prasad, Davide Rossi, Luca Benini, Francesco Conti
Presenter: Sergio Mazzola, ETH Zürich, Switzerland

Motivation

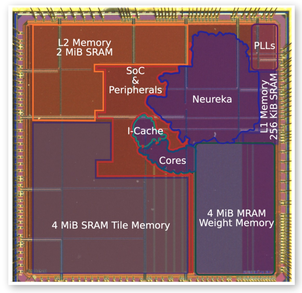
Shared-memory, tightly-coupled **heterogeneous cluster**
- Large local memories
- Diverse, powerful accelerators
- General-purpose control cores



Occamy [Paulin et al., VLSI Symp 2024]

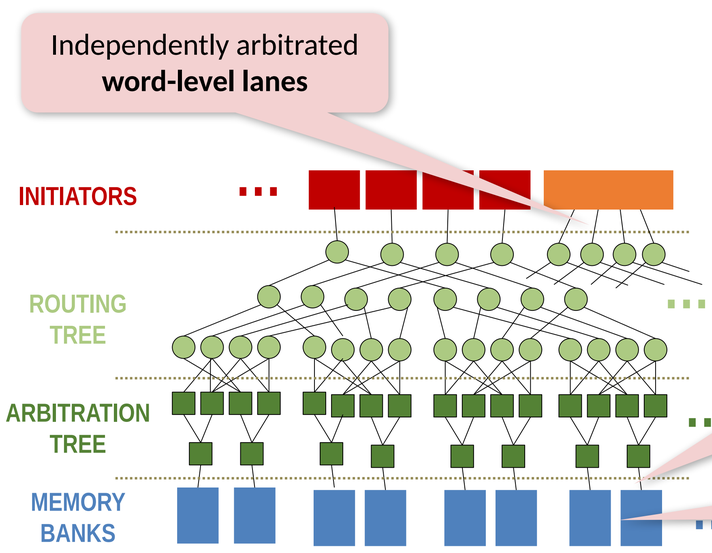


DataMaestro [Yi et al., DAC 2025]



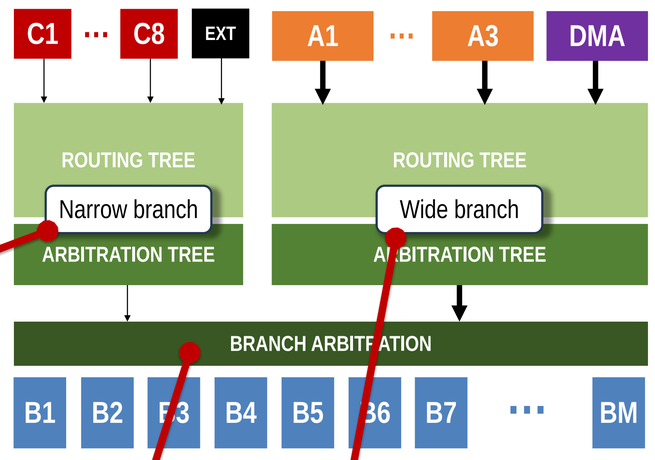
Siracusa 16nm [Prasad et al, JSSC 2024]

High performance, energy efficiency, flexibility
- End-to-end machine learning
- From edge to high-performance



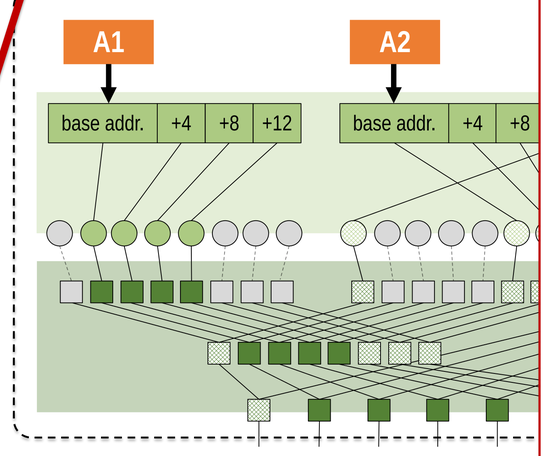
Architecture

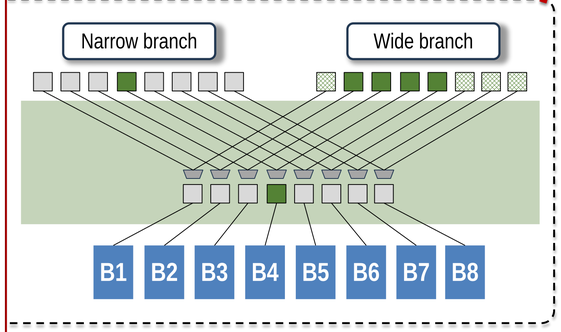
Heterogeneous Cluster Interconnect
- Handle **different traffic** independently
- Fine-grain, word-level access **only where needed**



Narrow branch
- Fully connected log xbar
- Word-granular, max flexibility

Wide branch
- Single access to contiguous addresses
- Handshaking shared across words, atomic grant



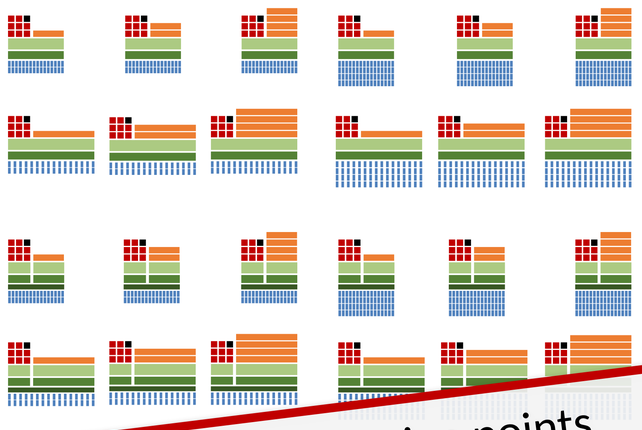


Branch arbitration
- Branch-level HoL blocking
- Runtime-configurable policy

github.com/pulp-platform/

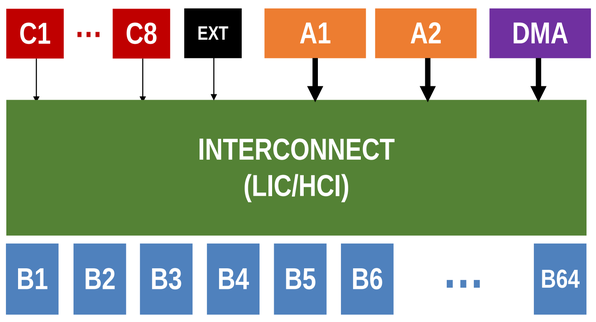
Experimental Procedures

Physical scalability analysis
- GlobalFoundries 22nm
- **24 interconnect configs**
  - **8** 32b cores + **1** ext
  - **1/2/4** **256/512b** accelerators
  - **32/64** TCDM banks
  - **LIC** vs **HCI**
- Corner: SS, 0.72V, - 40°C
- 500 MHz – 1.5 GHz
- 3 **Pareto-optimal** points per config for place&route



240 synthesis design points
72 place&route'd netlists!!!

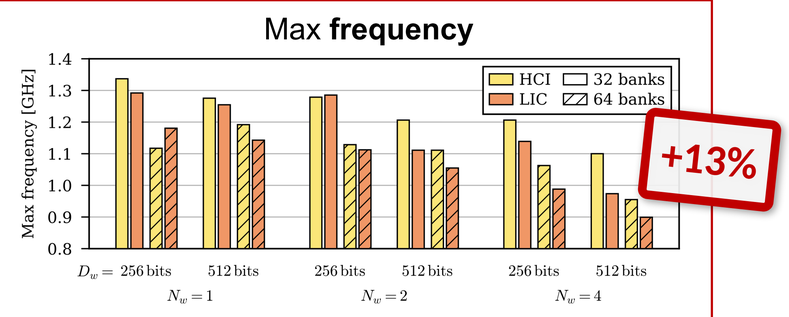
Performance benchmarking



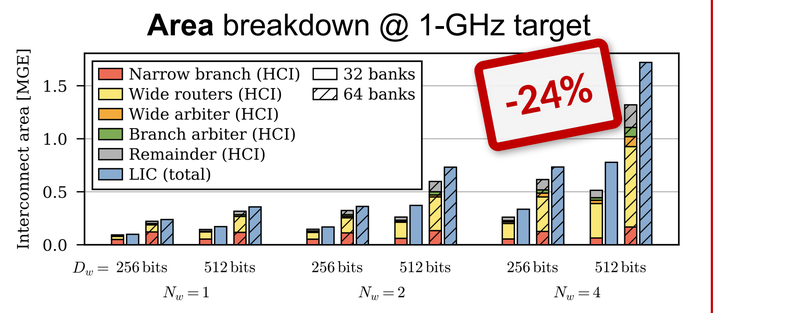
- 8 32-bit cores, 3 256-bit initiators
- 64-bank 256 KiB TCDM
- Workloads with diverse **narrow/wide** and **wide/wide traffic overlaps**

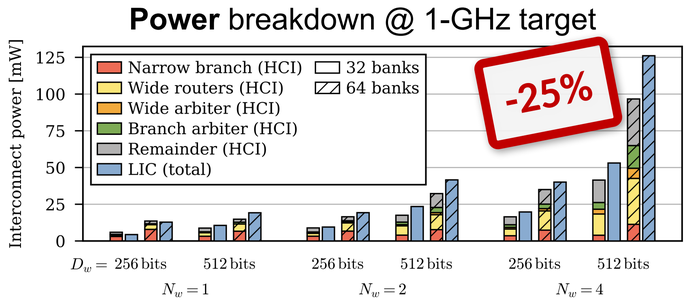
Experimental Results

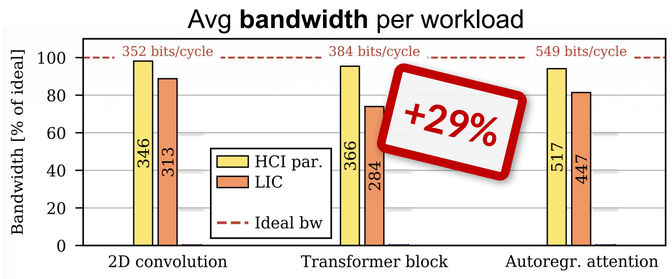
- Similar timing, area, power for **small configs**
- Critical path: dominated by SRAM read delay



- The more complex, the bigger the gap **HCI**/**LIC**
- **Wide branch** scales better
- Timing and area impact of **branch arbitration** still limited







- HCI energy efficiency: up to **+50%**
- Highest HCI improvement: high **arith. Intensity** + large **narrow/wide overlap**

ETH zürich
ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

PULP Platform
Open-Source Hardware, the way it should be!
pulp-platform.org