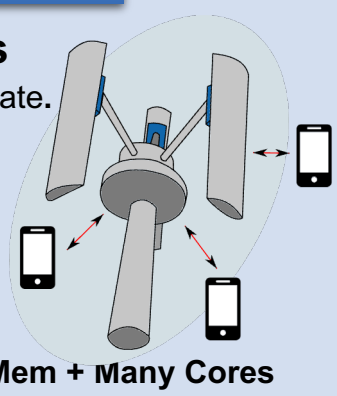
A 1024 RV-Cores Shared-L1 Cluster with HBM Link for Low-Latency 6G-SDR

Yichao Zhang[1], Marco Bertuletti[1], Chi Zhang[1], Samuel Riedel[1], Alessandro Vanelli-Coralli[1,2], Luca Benini[1,2]
[1]Integrated Systems Laboratory, ETH Zurich
[2]Department of Electrical, Electronic, and Information Engineering, University of Bologna

Introduction and Motivation

- LTE → 4G → 5G → **6G < 10 YEARs**
  - Network data traffic grows at Moore's rate.
- **Software Defined Radio:**
  - Reduce time to market;
  - Adaption to evolving standards;
  - Increase return on investments;
- **6G workload + Programmable = Big Mem + Many Cores**
  - Shared-L1 Mem reduces data split, merge & transfer;
  - Parallel execution + Final Sync;
  - Streamlined programming model.



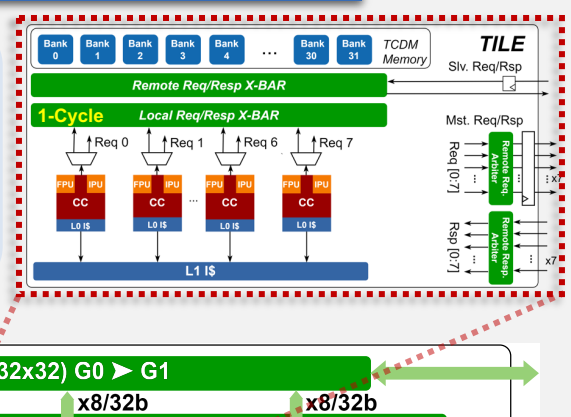
We take 1024 Cores Share L1 into ONE Cluster for 6G-SDR

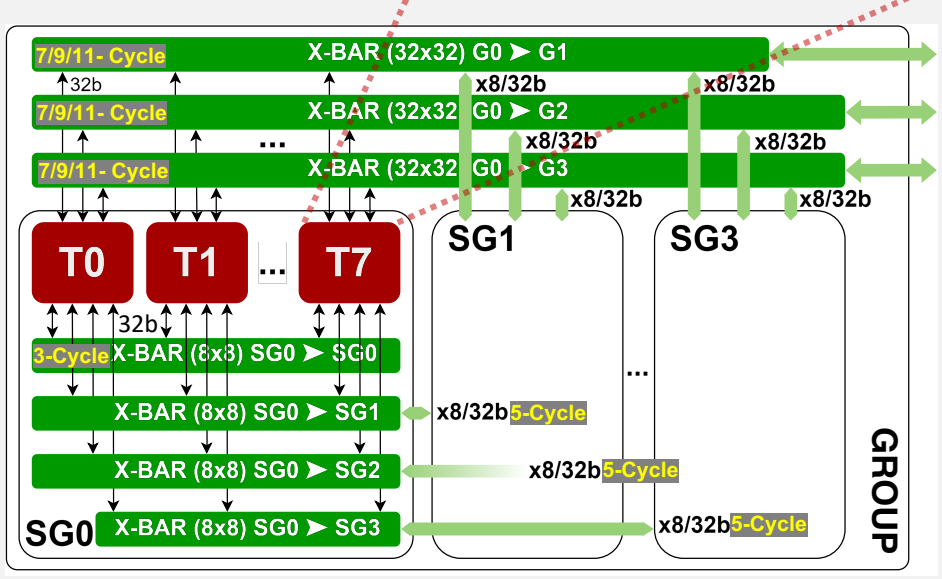
- Latency-tolerant, **programmable, NUMA**;
- **GF12nm**, Up to **924MHz** (TT/0.80V/25˚C);
- Shared **4MiB L1** (4096 Banks) SPM;
- **Only 7-11 Cycles** L1 Interconnection;
- Modular DMA + Hierarchical-AXI tree: **low-cost HBM Link.**

TeraPool

Low-Latency Cluster Architecture

Largest design so far in PULP: 1024 cores
RV32IMA-Xpulpimg
f32/f16 optional
4096 8b-MAC/cycle



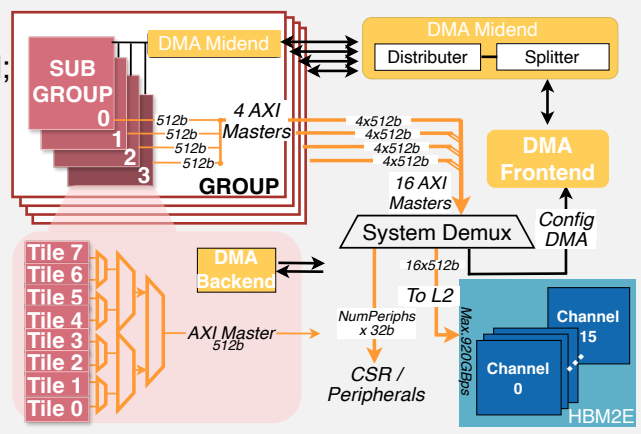


High Bandwidth Memory Link

Hierarchical L2-AXI:
To Main Memory;
Instruction Cache Refill;
DMA-Controlled.

Modular DMA:
Fe: configuration;
Me: transfer split;
Be: data mover;

Link to HBM2Es:
Two 16GiB Stacks;
AXI-L2 BW: 910GBps
(98% efficiency)

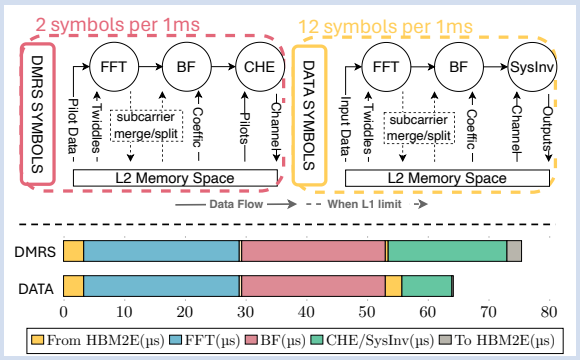


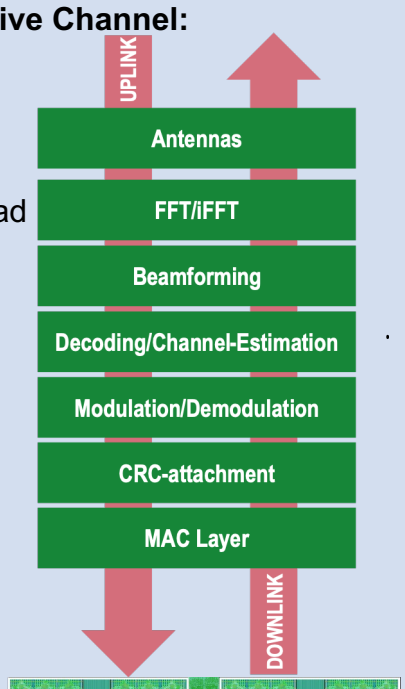
*Co-simulation with open-sourced cycle-accurate DRAMsys5.0 simulator

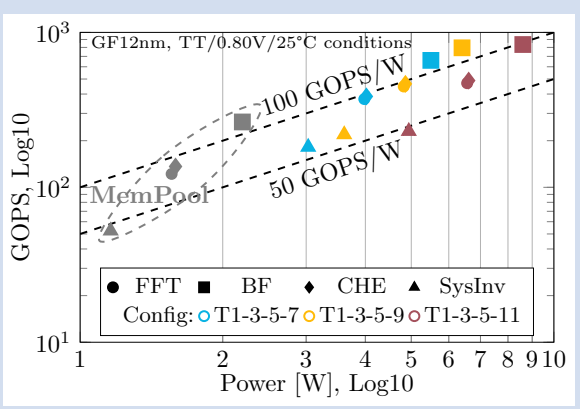
Energy-Efficient 6G SDR

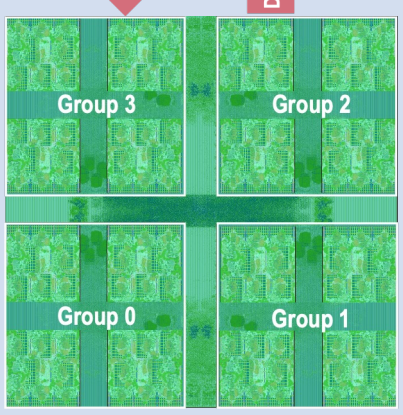
LOW-PHY: The Most compute Intensive Channel:

- 0.18 to **0.84 TOPS**
- Up to **125 GPOS/W**
- Less than **8.8W** cluster power
- Only **9% of data movement** overhead









Relative Work

1. *"TeraPool-SDR: An 1.89TOPS 1024 RV-Cores 4MiB Shared-L1 Cluster for Next-Generation Open-Source Software-Defined Radios," in GLSVLSI, 2024.*
2. *"Efficient parallelization of 5G-PUSCH on a scalable RISC-V many-core processor," in DATE, 2023.*
3. *"Mempool: A scalable manycore architecture with a low-latency shared l1 memory," IEEE Trans. Comput., 2023.*
4. *"Fast Shared-Memory Barrier Synchronization for a 1024Cores RISC-V Many-Core Cluster", in SAMOS XXIII, 2023.*

github.com/pulp-platform/mempool

pulp-platform.org
@pulp_platform
yiczhang@iis.ee.ethz.ch