A 10-core SoC with 20 Fine-Grain Power Domains for Energy-Proportional Data-Parallel Processing over a Wide Voltage and Temperature Range

Thomas Benz, Luca Bertaccini, Florian Zaruba, Fabian Schuiki
Frank K. Gürkaynak, Luca Benini
Integrated Systems Laboratory (IIS)
ETH Zürich, Switzerland
{tbenz, lbertaccini, zarubaf, fschuiki, kgf, lbenini}@iis.ee.ethz.ch

Outline

- Introduction
- Architecture
  - Overview
  - ISA Extensions
  - Power Gating
- Usecases
  - Datacenter
  - Extreme Edge
- Chip Results
- Comparison with SoA
- Conclusion

Introduction

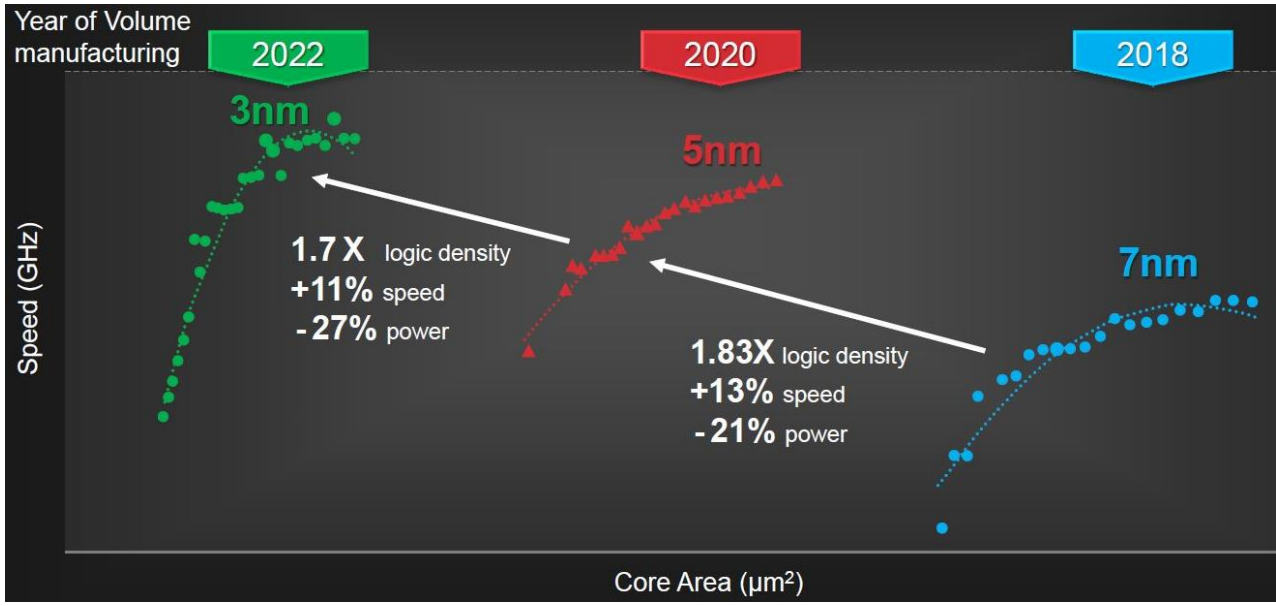
- Energy efficency: dominant factor for next-gen systems
  - **Scaling does not improve leakage**

- Leakage is an increasing problem
  - Increased silicon area + heterogenity
  - Higher integration density
  - ➢ High amount of **dark silicon**

- Leakage: no contribution to useful work
  - Wasted energy
  - Reduction of useful TDP

- A concern across the **compute continuum**
  - Datacenter
  - Extreme edge

Image: Unleashing the Future of Innovation/ISSC 2021, TSMC

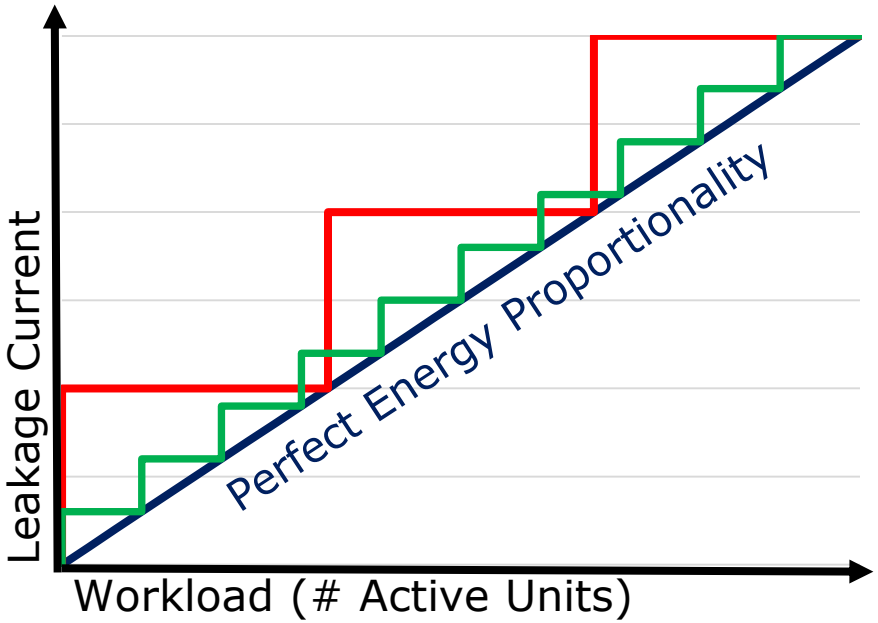


Equi-area chip:
power 1.24x up node-to-node!

Challenges & Contributions

Challenges

- **Fine-grain power gating** in manycore @ low overhead (area, timing) is **non-trivial**
- Usually: go for **coarse-grain** domains
  - ➢ limits **energy proportionality**

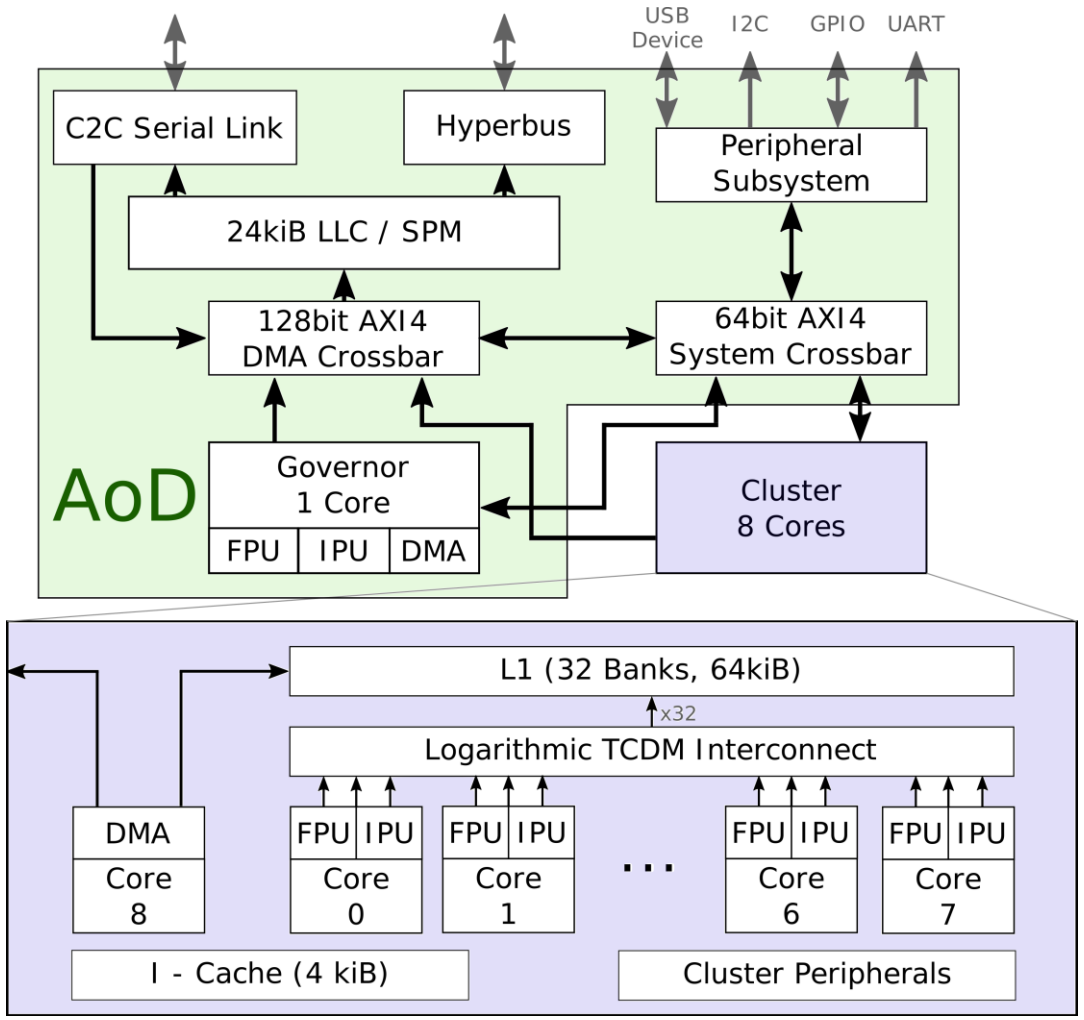


Contributions

- Manycore architecture with ISA extensions
  - Tuned for **energy efficiency**
  - Optimized **µ-architecture** to minimizes state
- First **fine-grain** power-manged RISC-V multi-core chip
  - Approach energy proportionality
  - Implemented and validated over a **wide freq. – temp. range**
  - Present **sub-10ns** wake-up time of power domains

Architecture Overview

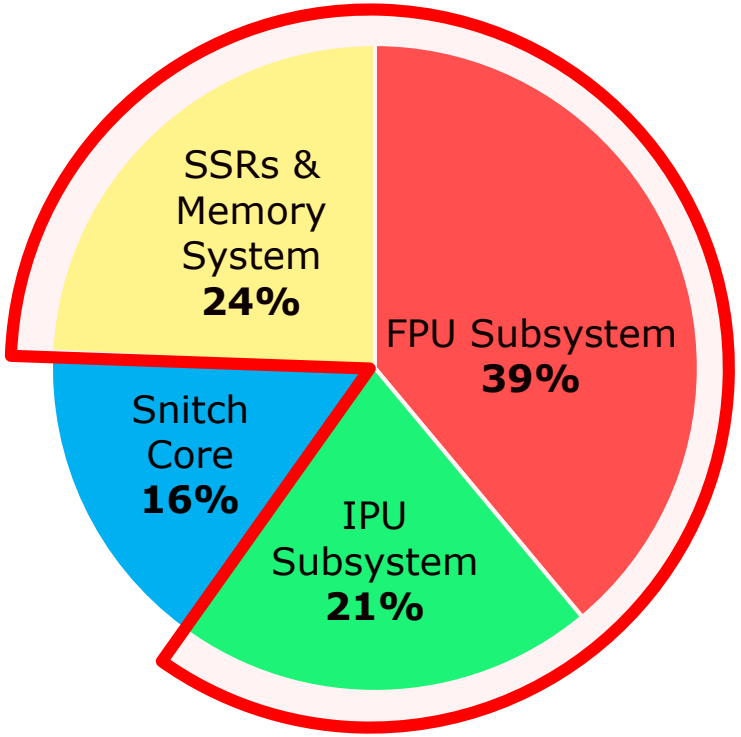
- Always on domain
  - Microcontroller
    - Governor
    - Management responsibilites
  - Peripherals
- Cluster
  - 8+1 Snitch Cores
    - Single-stage RISC-V
  - Core-Complex: core + FPU + IPU
  - 64kiB local scratchpad memory
  - ISA extensions
    - SSR
    - FREP
  - Fine-grain power gating

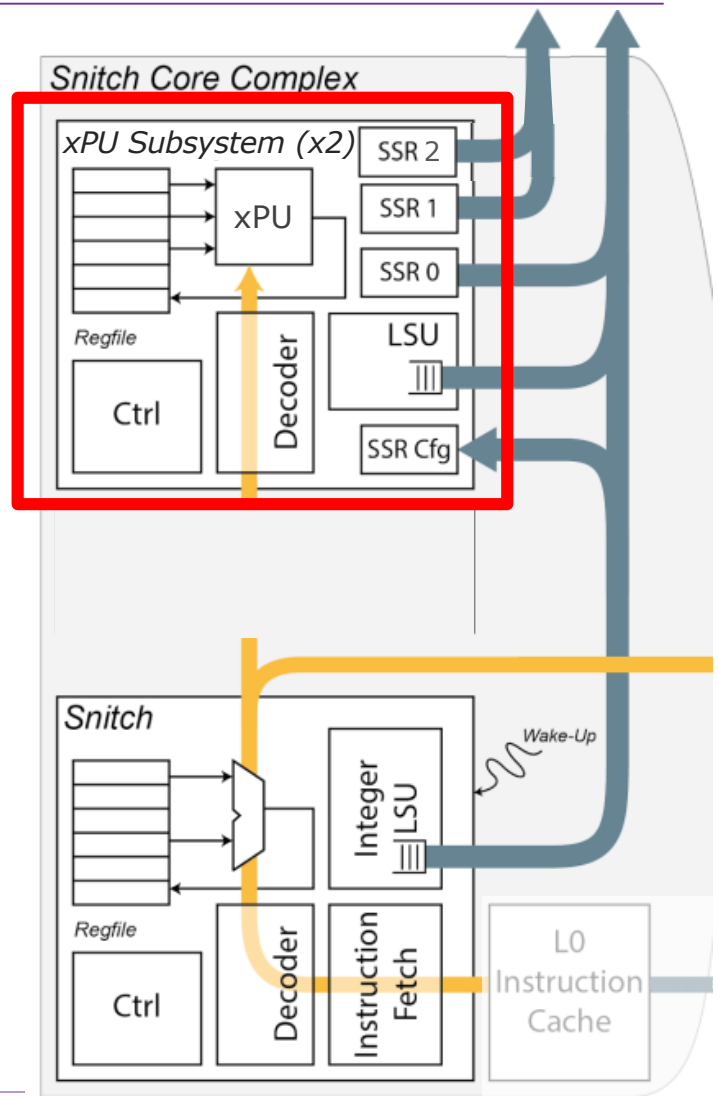


Snitch – highly power-manageable core

- Snitch Core Complex
  - Simple single-stage Snitch core
    - ***Control*** *core*
    - **Low** amount of logic
  - IPU, FPU functional units

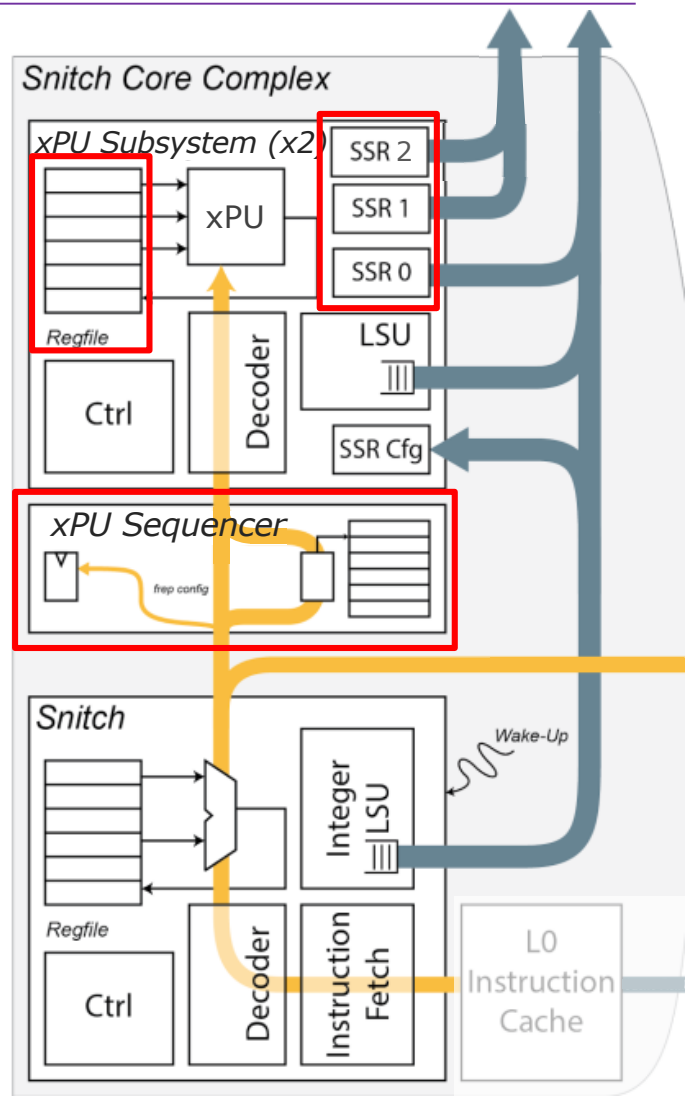
84% of the area of a Core Complex is power gated





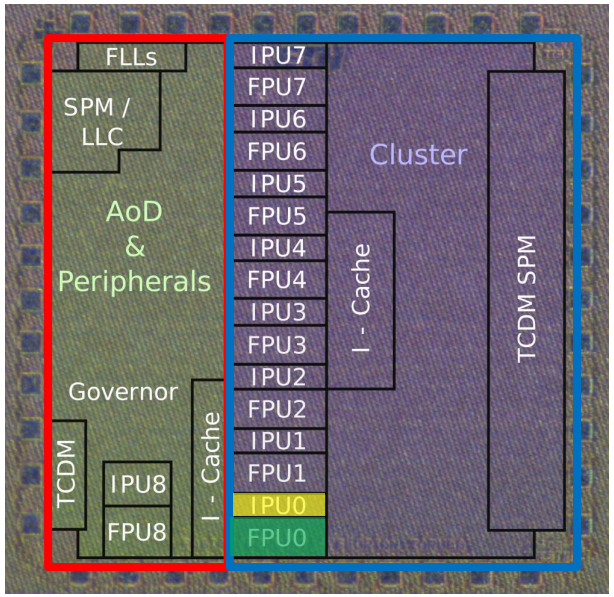
Snitch – highly power-manageable core

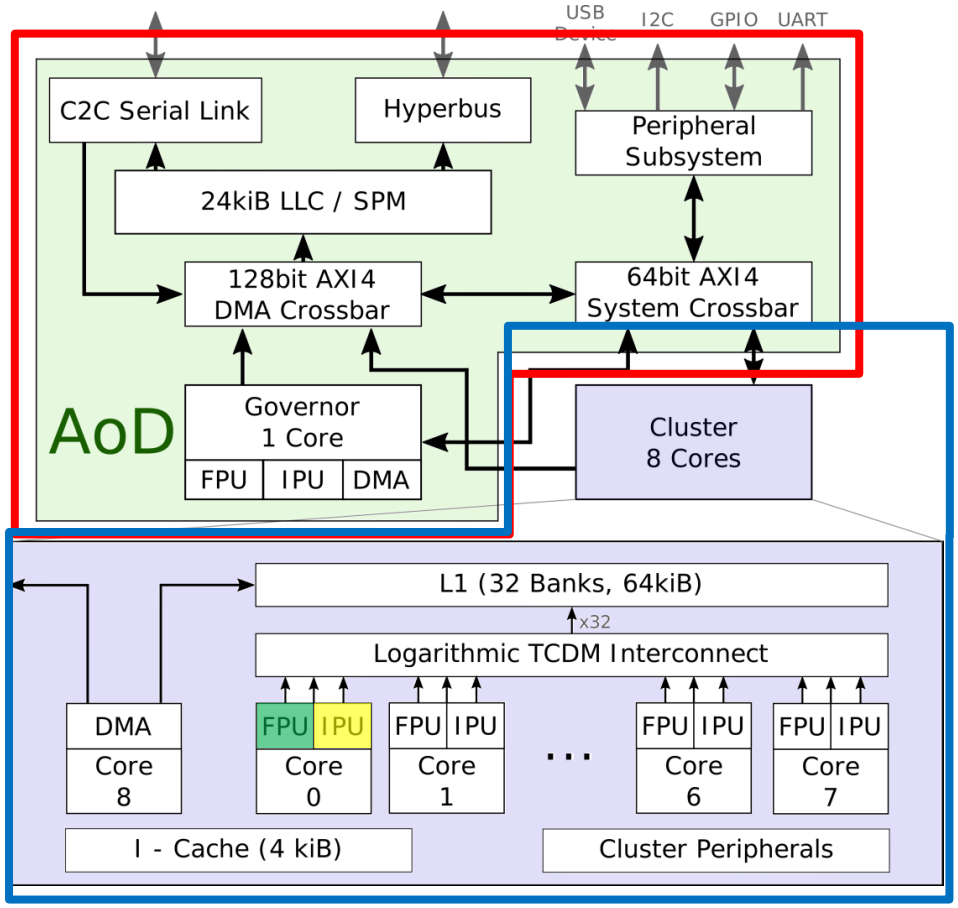
- Snitch Core Complex
  - Simple single-stage Snitch core
    - ***Control*** *core*
    - **Low** amount of logic
  - IPU, FPU functional units
- Goal: Utilize functional units **>80%**
  - **Superscalar out-of-order** cores
    - Large amount of **state & deep pipeline** stages in func. units
    - ➢ **Highly adverse to power management!**
  - Our solution: **simple** RISC-V ISA extensions
    - **SSR**: streaming semantic registers
    - **FREP**: hardware loop
    - Only **minimal** architectural state in power gated region
    - Algorithm using SSR, FREP: only **temporary** data
    - Optimized microarchitecture: only **4** pipeline stages



Power Gating Granularity

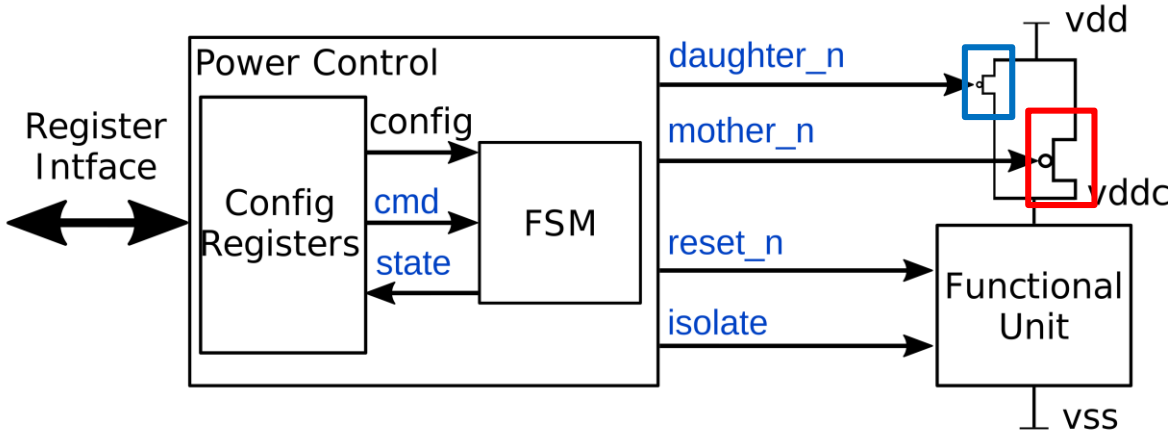
- AoD Domain
  - Governor Core
  - Not power gated
  - ➢ Focus on cluster
- Cluster domain
  - **Coarse-grain**
  - Entire Cluster
- Functional Units
  - **Fine-grain**
  - Individual FPU domains
  - Individual IPU domains





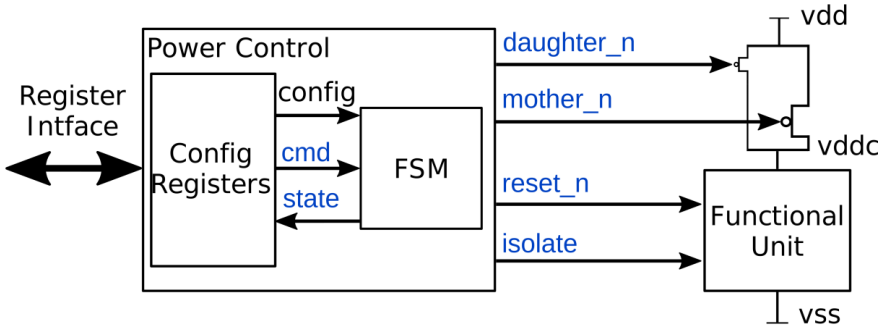
Power Control

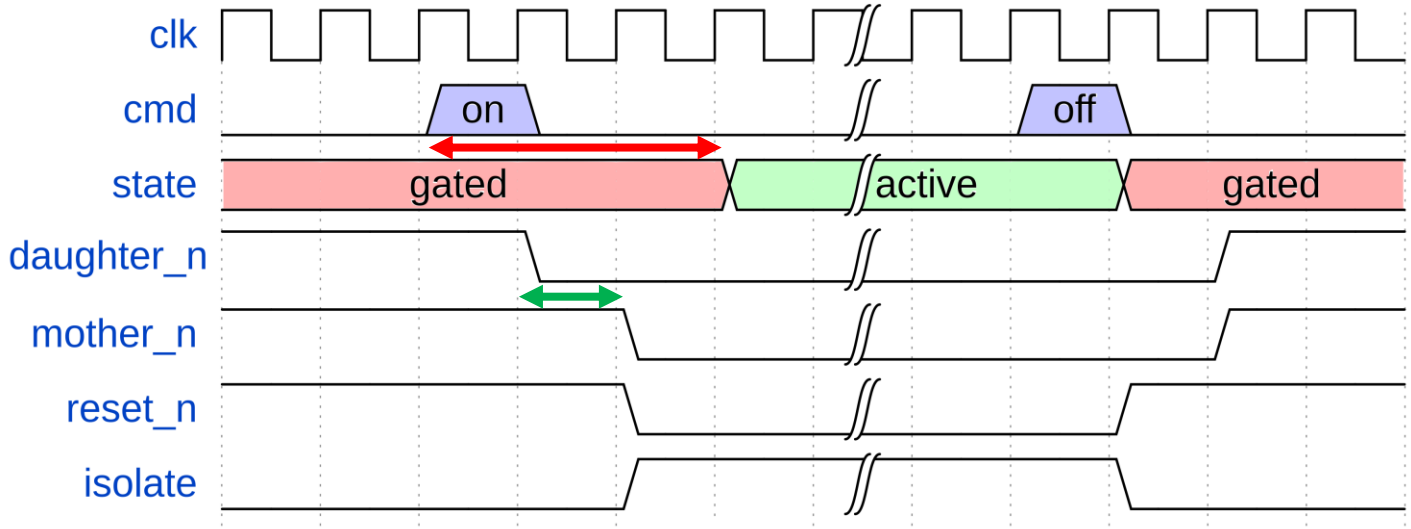
- Power control module
  - Memory-mapped register interface
  - Finite state machine
  - **Programmable** sequence
  - Only **11.4 kGE**
- Header power gates
  - Mother – Daughter
  - Reduce peak inrush current
- Isolation
  - Prevent hardware from injecting wrong transfers



Power Control Sequencing

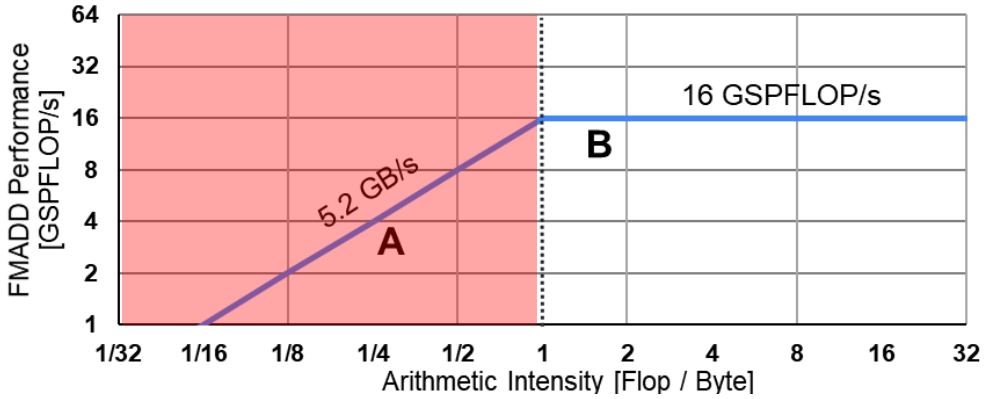
- 4-stage process
  - Timig configurable
  - **Single-write** power toggle
- Mother – Daugther delay
  - Most critical
  - Mitigate spike in inrush current
- Power toggle speed
  - 3 AoD clock cycles
  - ~6 cluster clock cycles
  - **Sub-10 ns**

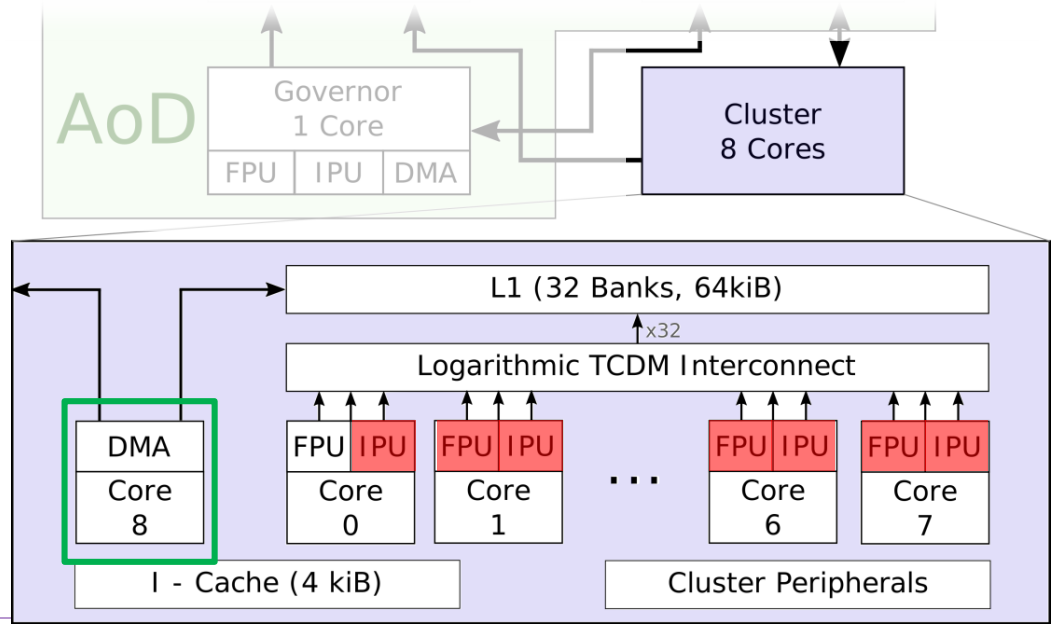




Use Case: Datacenter

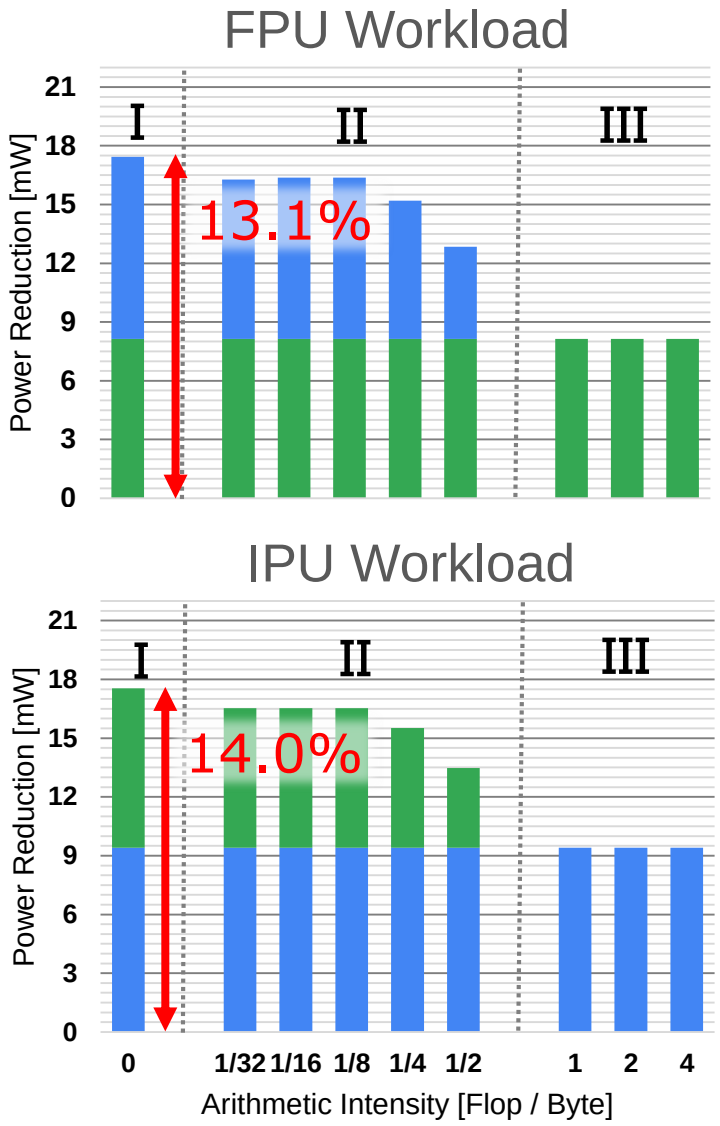
- Cluster-centric application
- 2 usecases for power gating:
  - IPU / FPU workloads are not mixed
    - ➢ Gate unused unit type
  - Memory-bound regimes
    - ➢ Gate stalling units
- Control is done by data movement core
  - Alreday used for data orchestration
  - Insigth in xPU utilization
  - Decentralized, scalable control





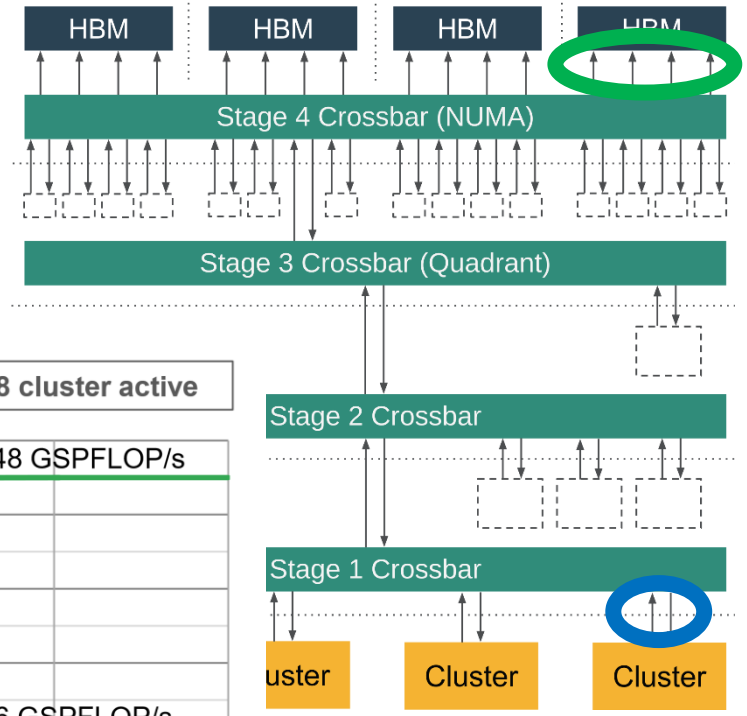
Use Case: Datacenter - Results

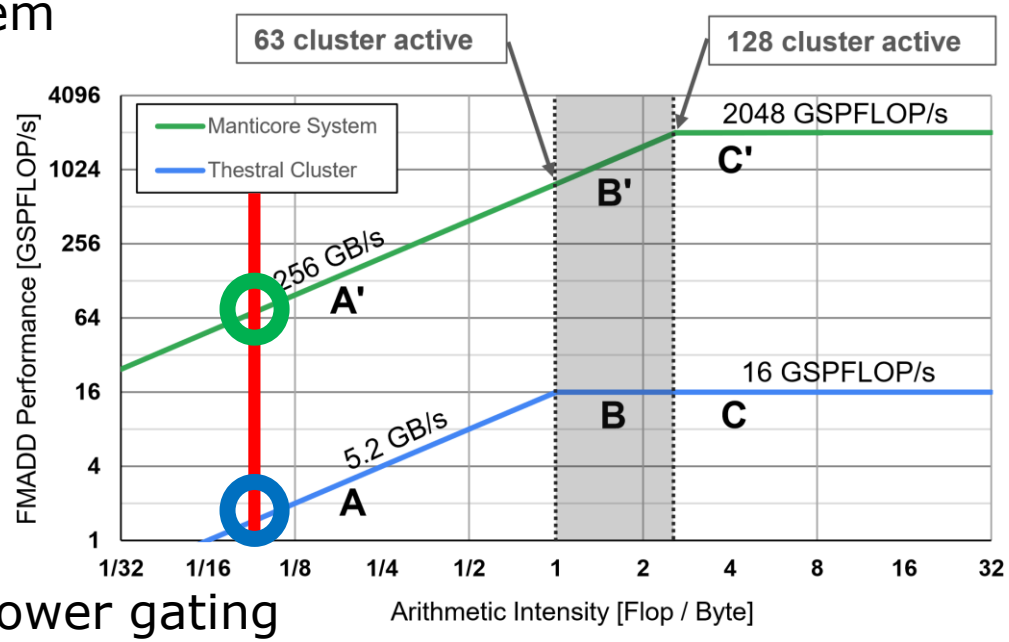
- 0.9V, 75°C, running at 850 MHz
- Running xPU workloads
  - Variable arithmetic intentity
- FPU Workloads
  - Power gate IPU units: 6.5% power reduction
  - Gate stalling units
    - In fully memory-bound region (I): up to 13.1%
- IPU Workloads
  - Power gate FPU units: 8.0% power reduction
  - Gate stalling units
    - In fully memory-bound region (I): up to 14.0%
    - Higher relative gain: IPUs consume less power



Study: Manticore

- Manticore Architecture*
  - **1024** cores on a chiplet
  - Organized in 128 clusters
  - **HBM2** memory interface
  - 4 chiplets: Manticore System
- AXPY FPU workload
  - Power gate all IPU units
  - 1/12 SP FLOP / Byte
  - **65 cluster fully gated**
  - 63 cluster: **1 FPU** active
- **41.4%** power reduction
  - **Coarse-** and **fine-grain** power gating
  - 15.18W to 8.96W



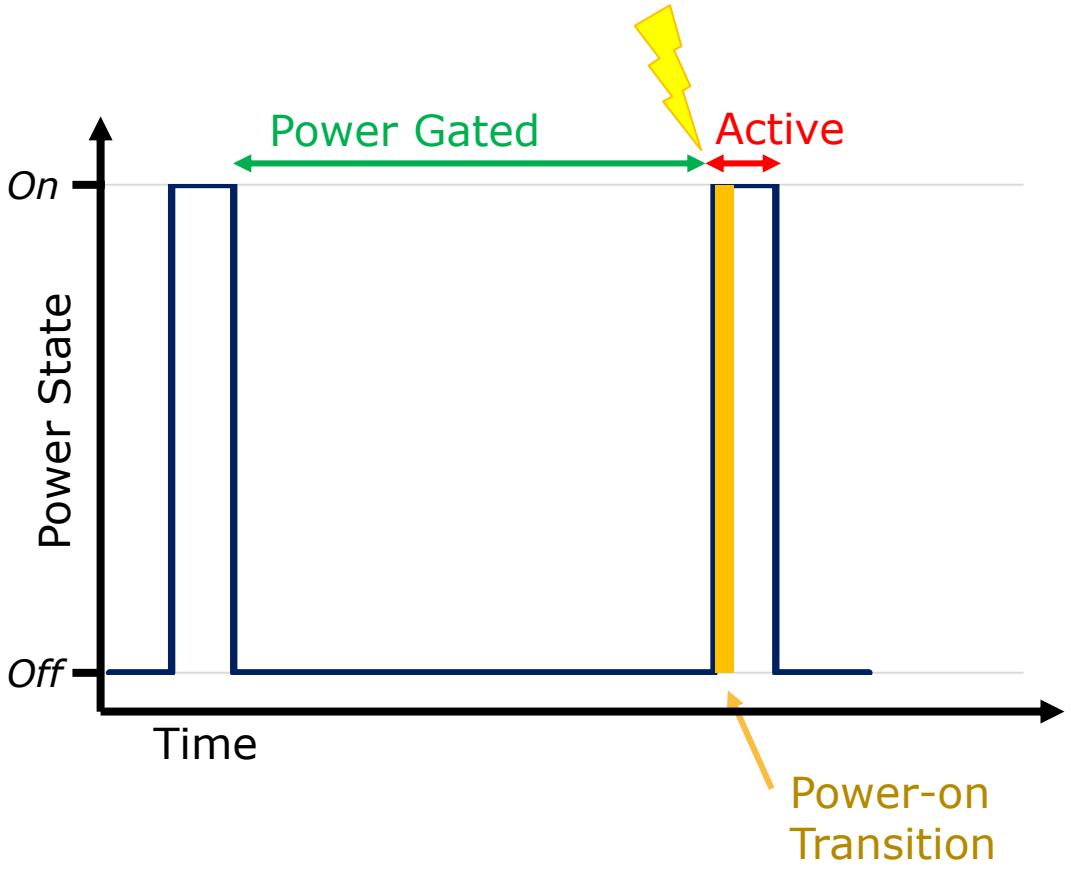


* F. Zaruba, F. Schuiki and L. Benini, "Manticore: A 4096-Core RISC-V Chiplet Architecture for Ultraefficient Floating-Point Computing," in *IEEE Micro*, vol. 41, no. 2, pp. 36-42, 1 March-April 2021

Use Case: Extreme Edge

- Microcontroller
  - Self-contained unit
  - 1 management core: AoD **Governor**
  - 8+1 core GP compute accelerator

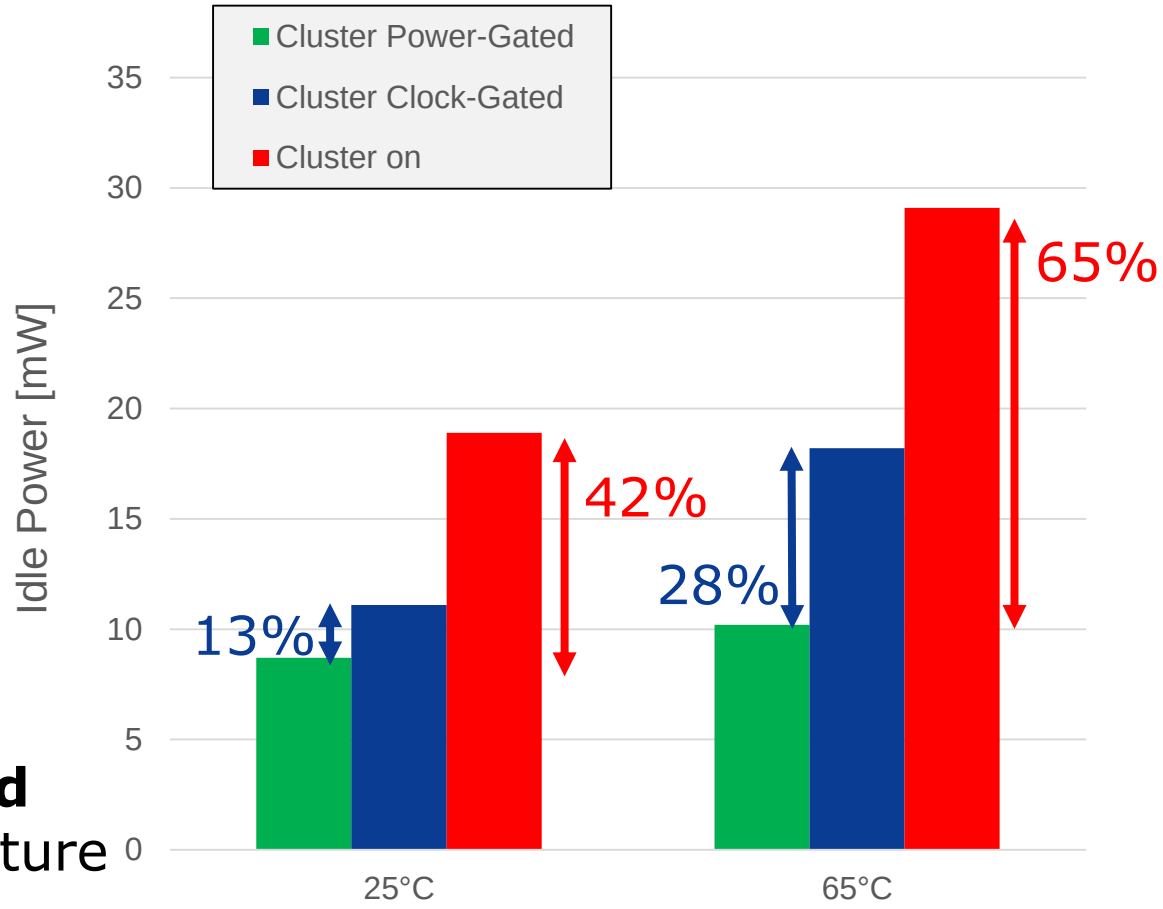
- Usecase for power gating
  - Sporadic need for **massive** compute
  - **Gate cluster fully** during **idle**
  - Need to reduce power-on transition time
  - Near-threshold for dynamic efficency
    - Worsens **ratio** of leakage
    - Slower clock: **less cycles** for transition



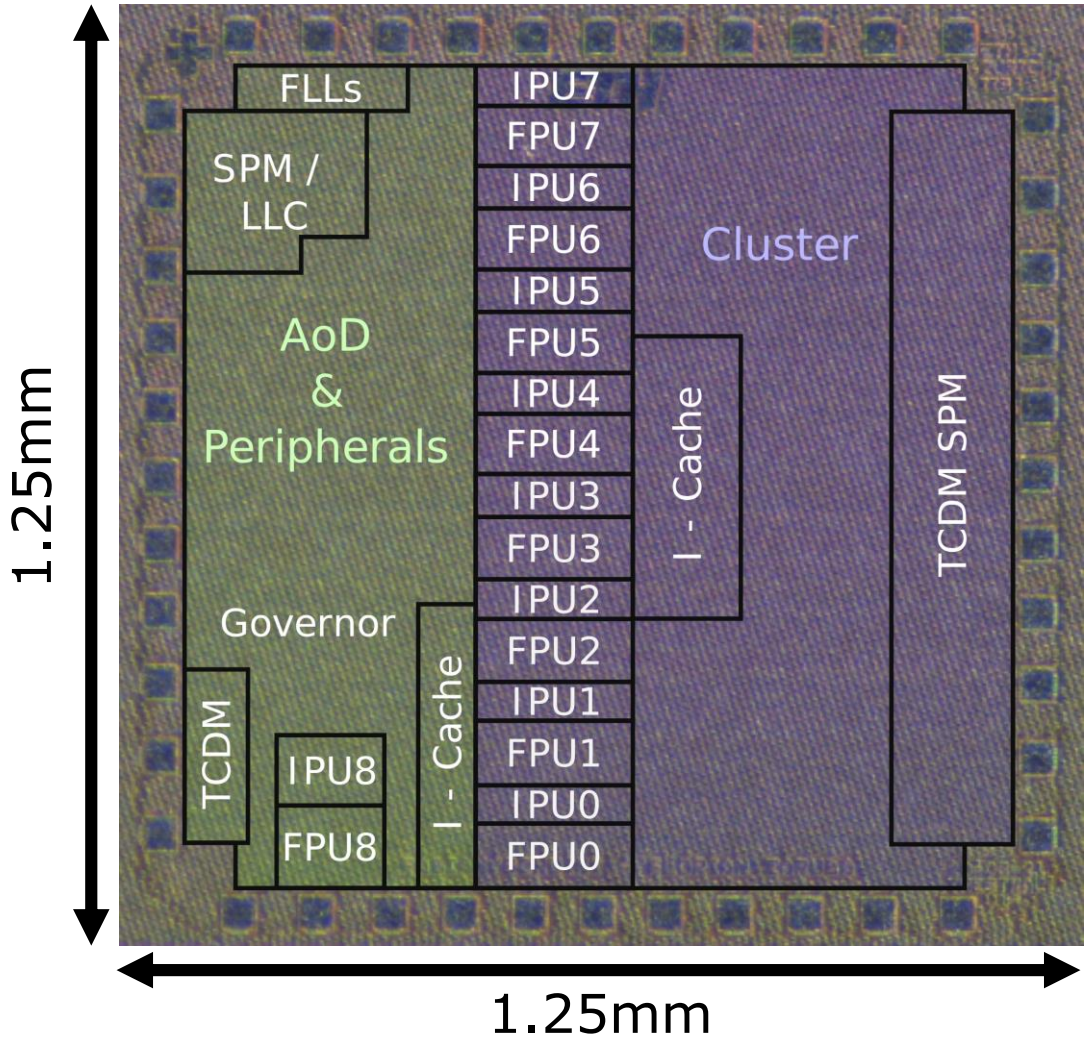
Use Case: Extreme Edge - Results

- Nominal operating point
  - Edge node
  - 0.6V, 25°C. Near-threshold
  - Up to 42% power reduction
- Extended temperature range
  - Automotive applications e.g.
  - 0.6V, 65°C
  - Up to 65% power reduction

➢ Near-threshold operation **remains energy efficient** also for **advanced nodes** and **high** operating temperature



Chip Results



Technology	GF 22nm FDSOI
Chip Area	1.56 mm^2
VDD Range	0.6V – 0.9V
Memory size	64kiB L1, 24kiB L2
Logic Transistors	6 MGE
Frequency Range	32kHZ – 950MHz
# Controllable Power Domains	18 fine-, 1 coarse-grain

Power Domain	**Area**
Always on Domain	1.52 MGE
Cluster	2.69 MGE
FPU	57.0 – 58.7 kGE
IPU	31.3 – 32.2 kGE

Comparison With SoA

	SamurAI [2]	Vega [3]	Thestral (This work)	[6]	A64FX [5]
Application	IoT	IoT	IoT/HPC	HPC	HPC
Technology	28nm FSOI	22nm FDSOI	22nm FDSOI	28nm FDSOI	7nm FinFET
Die Size	4.5 mm^2	12 mm^2	1.56 mm^2	7.84 mm^2	-
Cores/ISA	32b Async RISC 32b / RISC-V	1 + 9 cores / 32b RISC-V	1 + 9 cores / 32b RISC-V	2 cores / 64b RISC-V	48 + 4 cores / Armv8-A
Accelerator	ML /Cypto	4 shared FPUs / HWCE	8 FPUs / 8 IPUs	Hwacha Vector	SVE / 512 bit SIMD
Maximum Frequency	350 MHz	450MHz	910 MHz	475 MHz	2.2 GHz
Voltage Range	0.45V - 0.9V	0.5V - 0.8V	0.6V - 0.9V	0.55V - 1.1V	-
On-chip SRAM (State Retention)	464kB	128kB (L1) / 1600kB s.r. (L2)	64kB (L1) / 24kB s.r. (L2)	256kB	32MB (L2 Cache)
Power Management	Clock & Power Gating	Clock &Power Gating	Clock & Power Gating	DVFS & Body-Biasing	Power Gating & DVFS
Granularity	On-Demand Unit	Cluster / SoC / Memories	Cluster / IPUs / FPUs	Tile (Core)	FPU Lane
Best INT Performance	1.5 GOPS	15.6 GOPS (8-bit)	6.8 GOPs (32-bit)	Not Available	Not Available
Best FP Performance	-	2 GFLOPS (FP32)	13.6 GFLOPS (FP32)	Not Available	3.4 TFLOPS (FP64)
Reaction Time	207ns	not available	10ns	<2us	few ms
Energy Efficiency	230 GOPS/W @ 110 MOPS (int8)	79 GFLOPS/W @ 1 GFLOPS	118 GFLOPS/W @ 7.2 GFLOPS	19.6 GFLOPS/W	16.9 GFLOPS/W (FP64)

Conclusion

- Thestral: **10**-core chip with **20** power domains
  - RISC-V manycore chip
  - Custom, light-weight **ISA extensions**
  - **Agile, sub-10 ns, aggressive, fine-grain** power management
- Wide range of applications
  - **IoT**: microcontroller with GP compute acceleration
  - **Datacenter**: compute cluster with decentralized, fine-grain power control
- Datacenter: up to **41.4%** power reduction in a 1024-core chiplet
- IoT: Up to **65%** power reduction near-threshold, **high** environmental temperature
- High energy efficiency of **118 GFLOPS/W** @ 7.2 GFLOPS