An Asynchronous Array of Simple Processors for DSP Applications

Zhiyi Yu, Michael Meeuwsen, Ryan Apperson,
Omar Sattari, Michael Lai, Jeremy Webb, Eric Work,
Tinoosh Mohsenin, Mandeep Singh, Bevan Baas

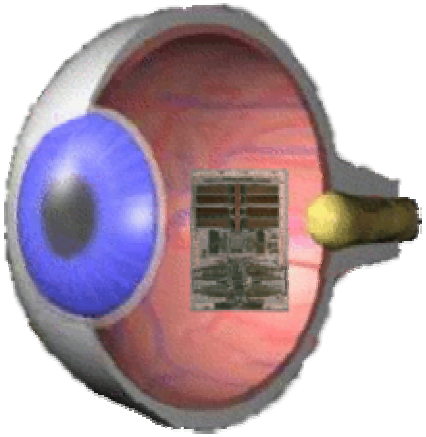
VLSI Computation Lab, ECE Department
University of California, Davis, USA

Outline

- **Project objectives**
- Key features of the AsAP processor
- Design of the AsAP processor
- Results

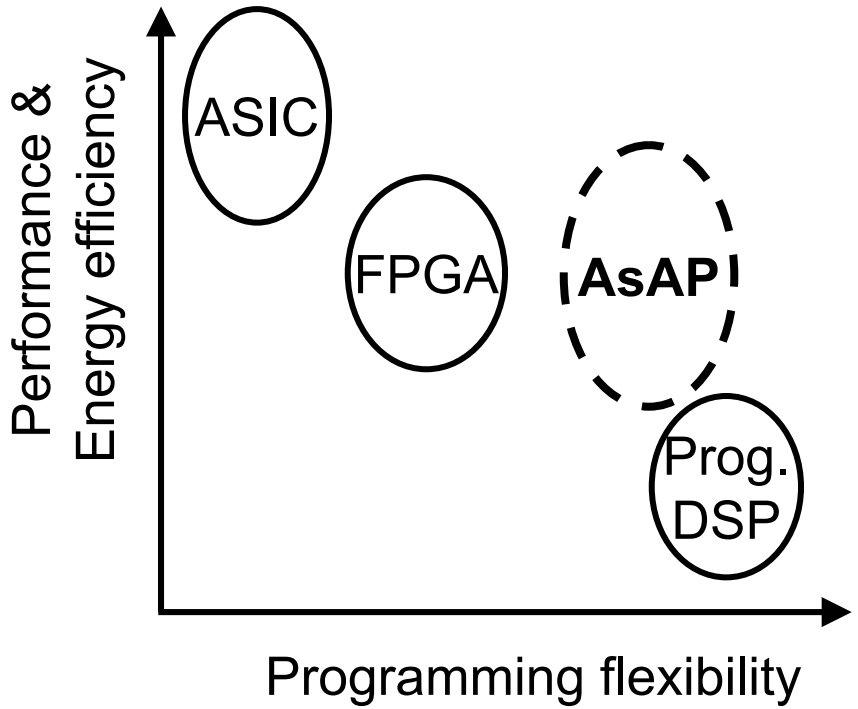
Target Applications

- Computationally intensive DSP and scientific apps
  - Key components in many systems
  - Require high performance
  - Limited power budgets
  - Require innovations in architecture and circuit design



Project Objectives

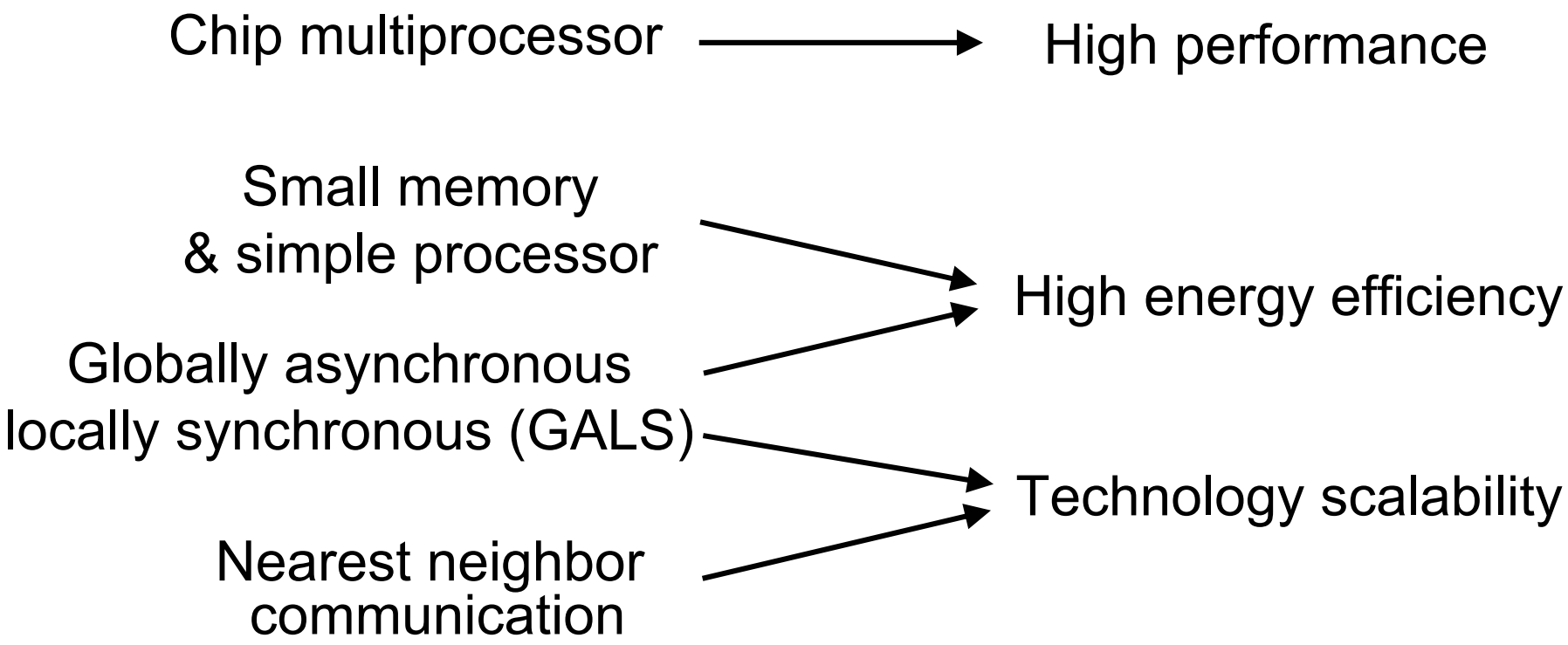
- Programming flexibility
- High performance
  - Throughput
  - Latency
- High energy efficiency
- Suitable for future fabrication technologies



Outline

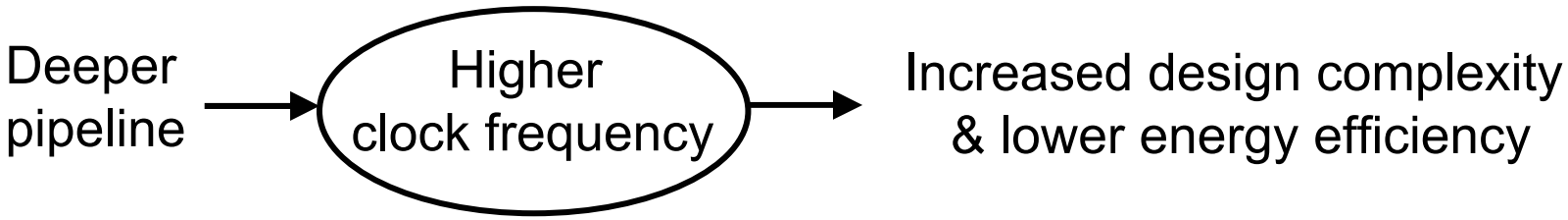
- Project objectives
- **Key features of the AsAP processor**
- Design of the AsAP processor
- Results

Key Features of the AsAP Processor



High Performance Through Chip Multiprocessor

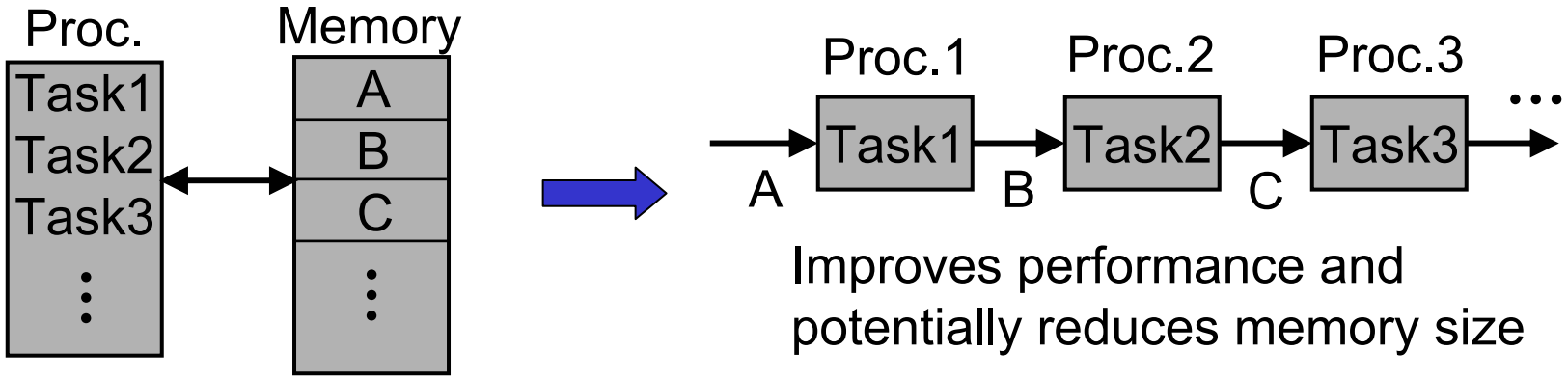
- Increasing the clock frequency is challenging



- Parallelism is more promising
  - Instruction level parallelism (VLIW, Superscalar)
  - Data level parallelism (SIMD)
  - Task level parallelism

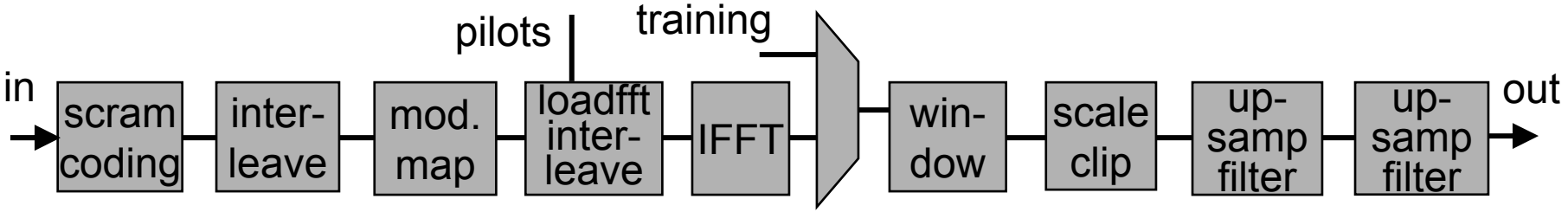
Task Level Parallelism

- Well suited for DSP applications



Improves performance and potentially reduces memory size

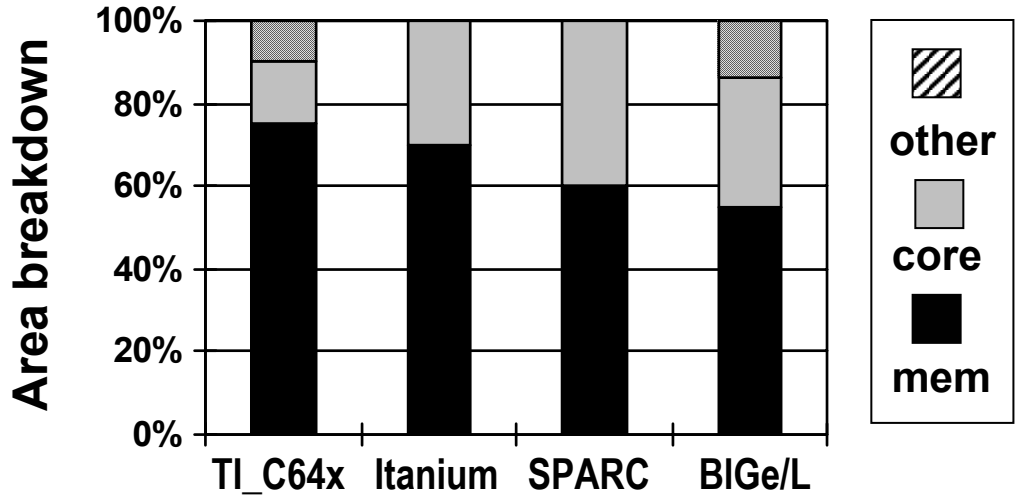
- Widely available in many DSP applications



802.11a/g wireless LAN (54 Mbps, 5 GHz) baseband transmit path

Memory Size in Modern Processors

- Memory occupies much of the area in modern processors — this reduces area available for the core and consumes large amounts of power
- Area and energy dissipation can be dramatically decreased with smaller memories



[ISSCC 02, 05]

Small Memory Requirements for DSP Tasks

- The memory required for common DSP tasks is quite small
- Several hundred words of memory are sufficient for many DSP tasks

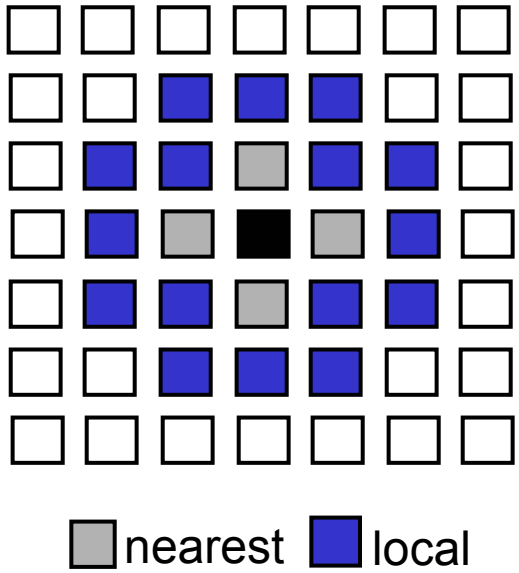
Task	IMem (words)	DMem (words)
N-pt FIR	6	2*N*
8-pt DCT	40	16
8x8 2-D DCT	154	72
Conv. coding (*k* = 7)	29	14
Huffman encode	200	330
N-pt convolution	29	2*N*
64-pt complex FFT	97	192
Bubble sort	20	1
N merge sort	50	*N*
Square root	62	15
Exponential	108	32

GALS Clocking Style

- The challenge of globally synchronous systems
  - Design difficulty when using high clock frequencies, long clock wires caused by large chip sizes, and large circuit parameter variations
  - High clock power consumption and lack of flexibility to independently control clock frequencies
- How about totally asynchronous design style
  - Lack of EDA tool support
  - Design complexity and circuit overhead
- The GALS compromise
  - Synchronous blocks each operating in their own independent clock domain

Wires and On Chip Communication

- Global wires are a concern
  - Their length doesn't shrink with technology scaling—assuming the chip size remains the same
  - The ratio of global wire delay to gate delay doubles each generation
- Methods to avoid global wires
  - Network on chip (NOC)
  - Local communication
  - Nearest neighbor communication

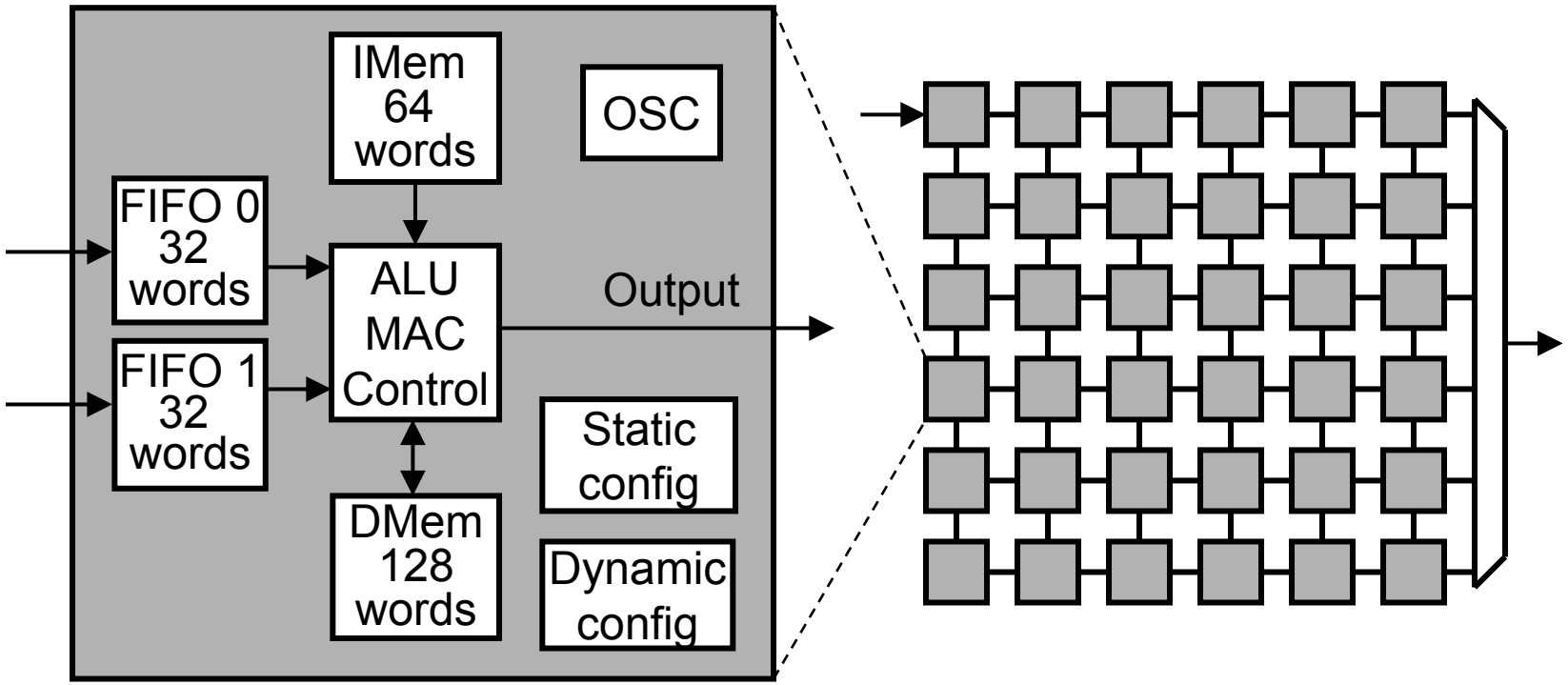


Outline

- Project objectives
- Key features of the AsAP processor
- **Design of the AsAP processor**
- Results

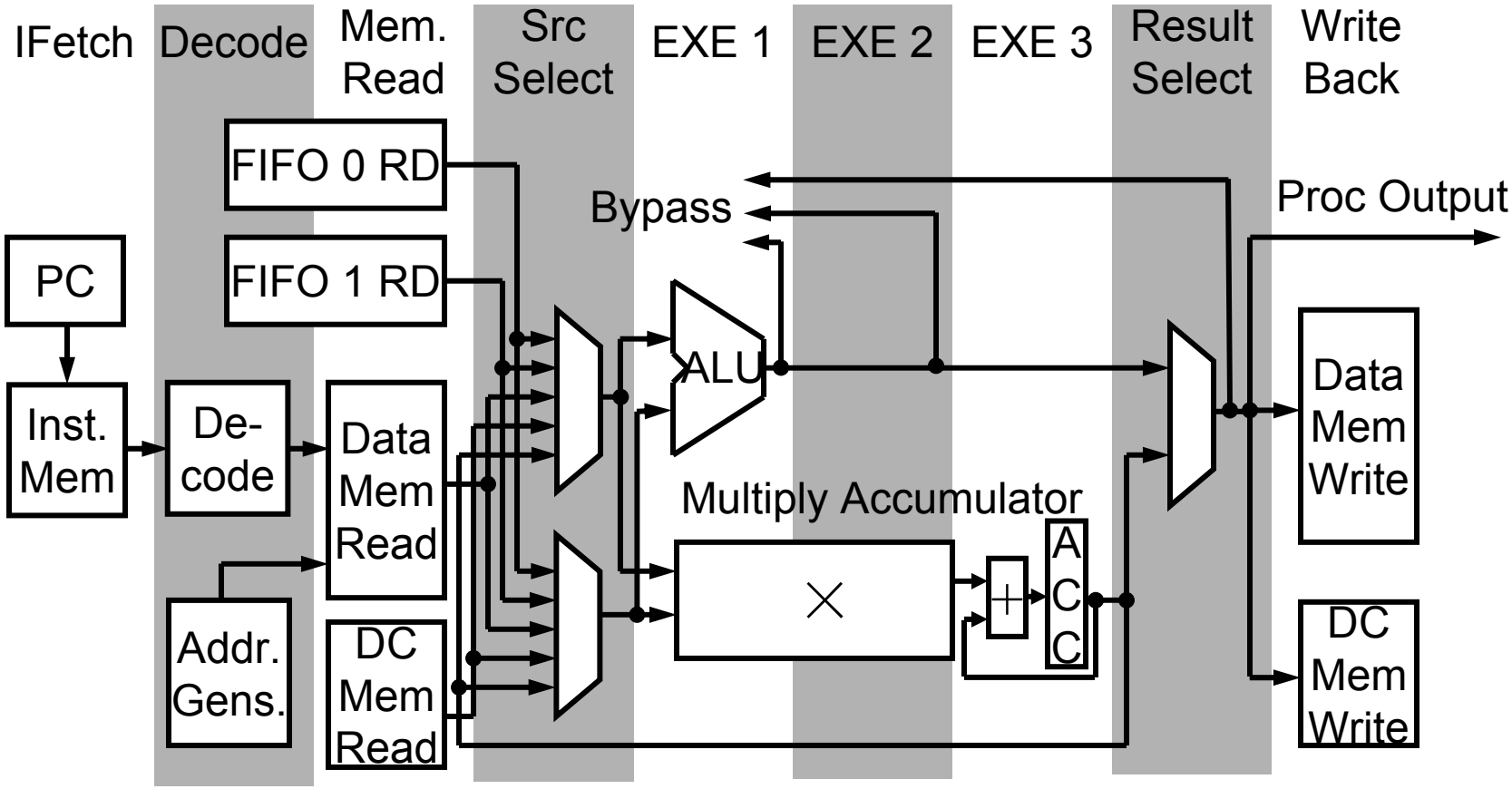
AsAP Block Diagram

- GALS array of identical processors
  - Each processor is a reduced complexity programmable DSP with small memories
  - Each processor can receive data from any two neighbors and send data to any of its four neighbors



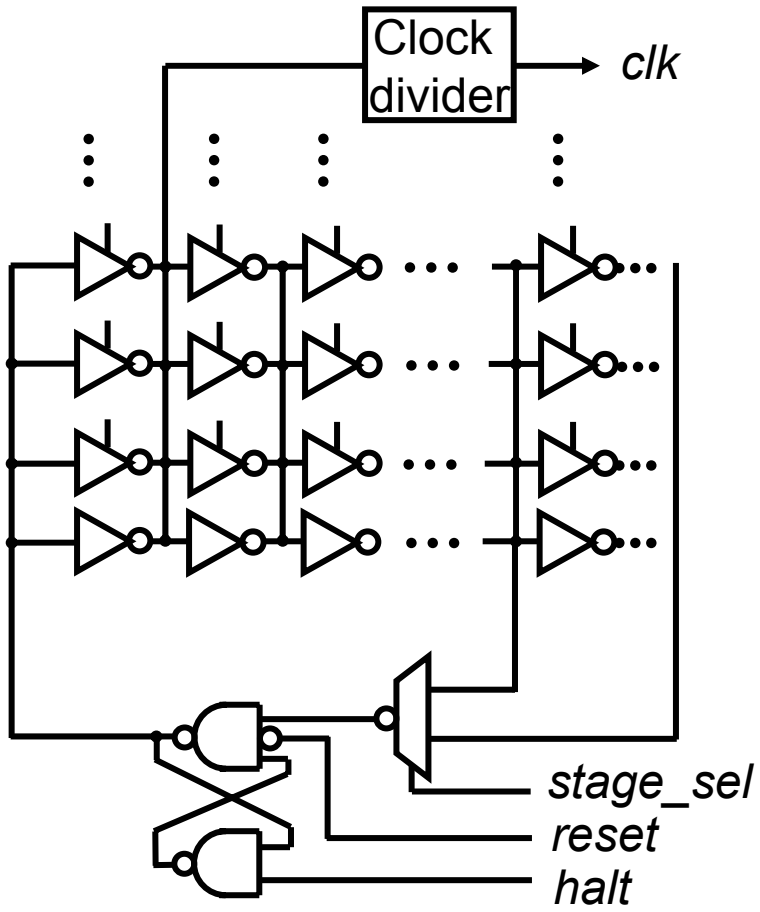
Single Processor Architecture

- 54 32-bit instructions, among which only Bit-Reverse is algorithm specific
- 9-stage pipeline



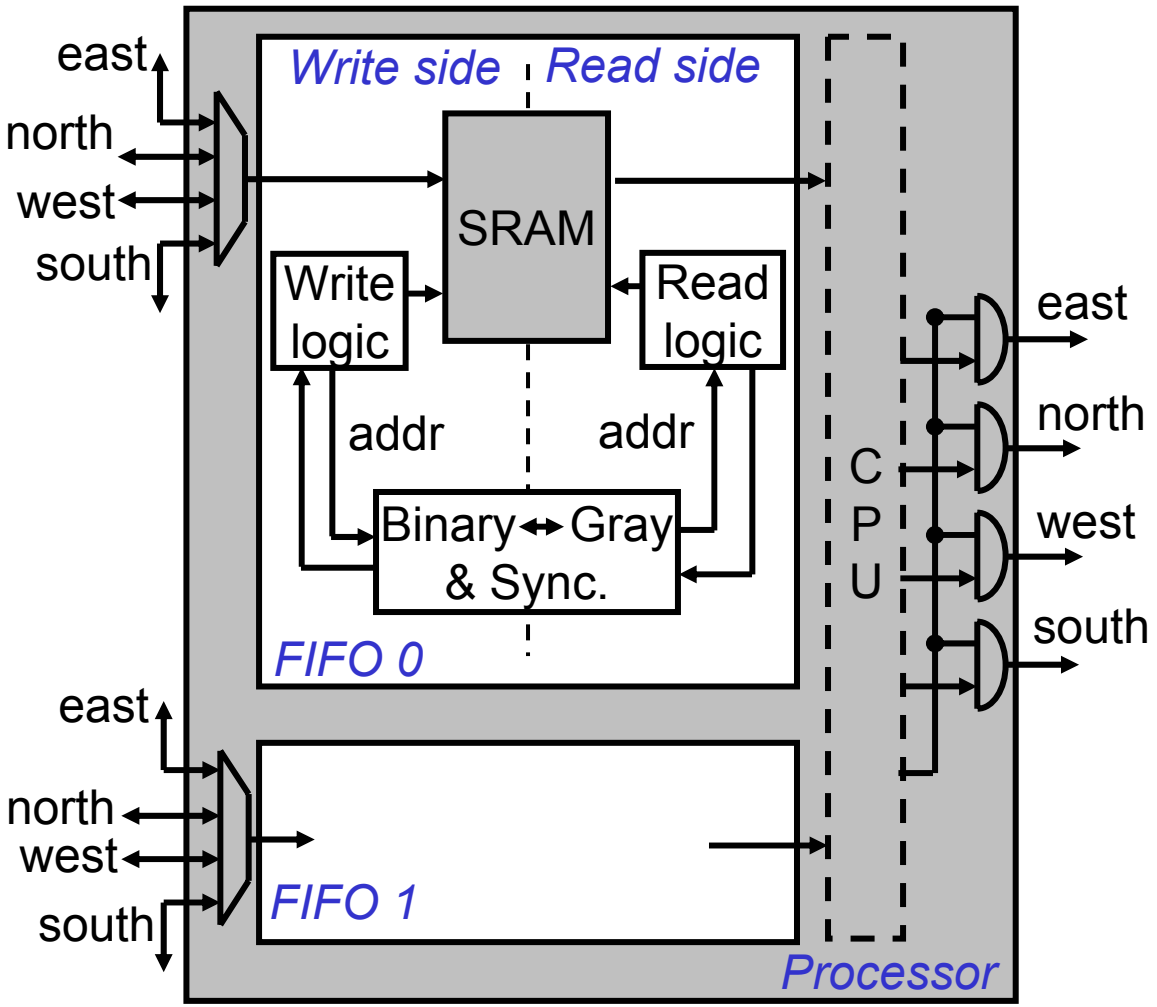
Programmable Clock Oscillator

- Standard cell based
- Configurable frequency
  - Delay tunable stages using 7 parallel tri-state inverters
  - 5 or 9 stage selection
  - 1 to 128 clock divider
- SR latch logic enables clean OSC halt
- Results
  - 1.66 MHz – 702 MHz
  - Max gap: 0.08 MHz (1.66 – 500 MHz)



Inter-processor Communication

- Each processor contains two dual-clock FIFOs
- Rd/Wr in separate clock domains
  - Gray coded Rd/Wr address across clock domains
- No FIFO failures in several weeks of testing with multiple procs



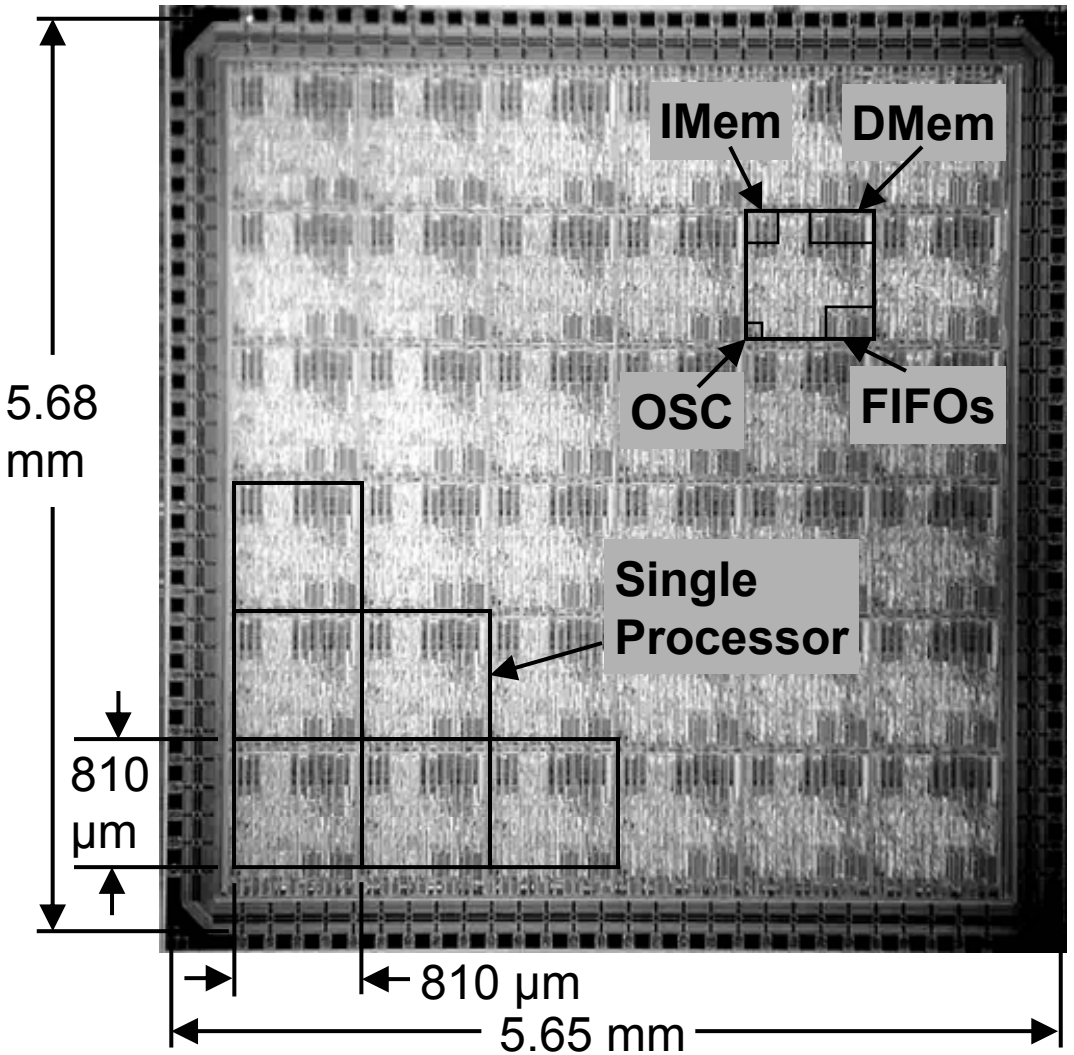
Advantages of the AsAP Clocking System

- Simplified clock tree design
  - The maximum span is < 1 mm in 0.18 μm
- Scalable – easy to add processors
- Improved energy efficiency
  - Clock halts in 9 cycles (processor dissipates leakage power only) and restores in < 1 cycle according to work availability
  - 53% and 65% power savings for two applications
  - Independent clock and voltage scaling (individual processor voltage scaling not implemented in this version)

Physical Design

- 0.18 μm TSMC
- Standard cell
- Design flow
  - Completely synthesized, except oscillator
  - Macro memory blocks used for four main memories
  - Completely auto placed and routed
- To simplify physical design, clock gating not implemented in this version

Chip Micrograph of the 6 x 6 Array



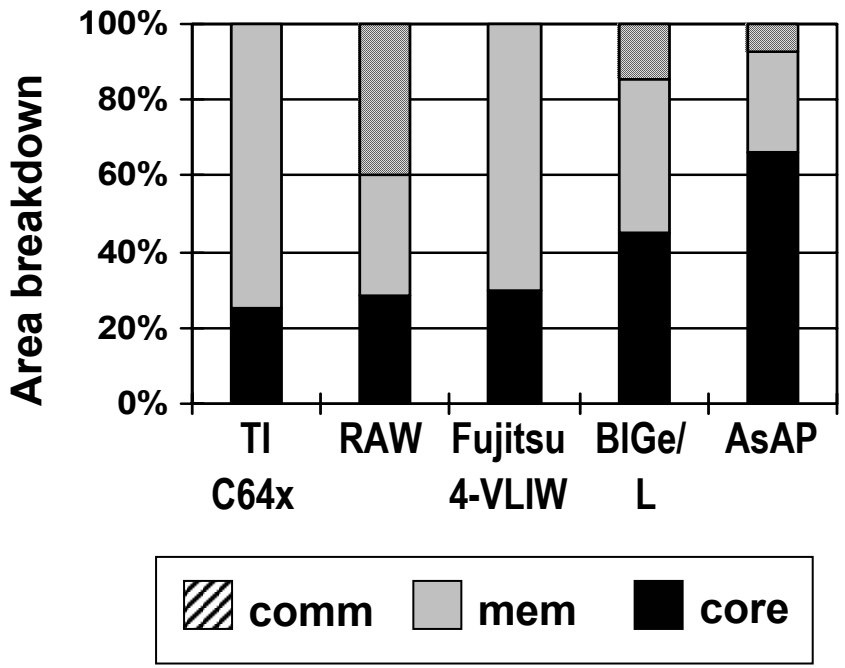
Transistors:	
1 Proc	230,000
Chip	8.5 million
Max speed:	475 MHz @ 1.8 V
Area:	
1 Proc	0.66 mm²
Chip	32.1 mm²
Power (1 Proc @ 1.8V, 475 MHz):	
Typical application	32 mW
Typical 100% active	84 mW
Worst case	144 mW
Power (1 Proc @ 0.9V, 116 MHz):	
Typical application	2.4 mW

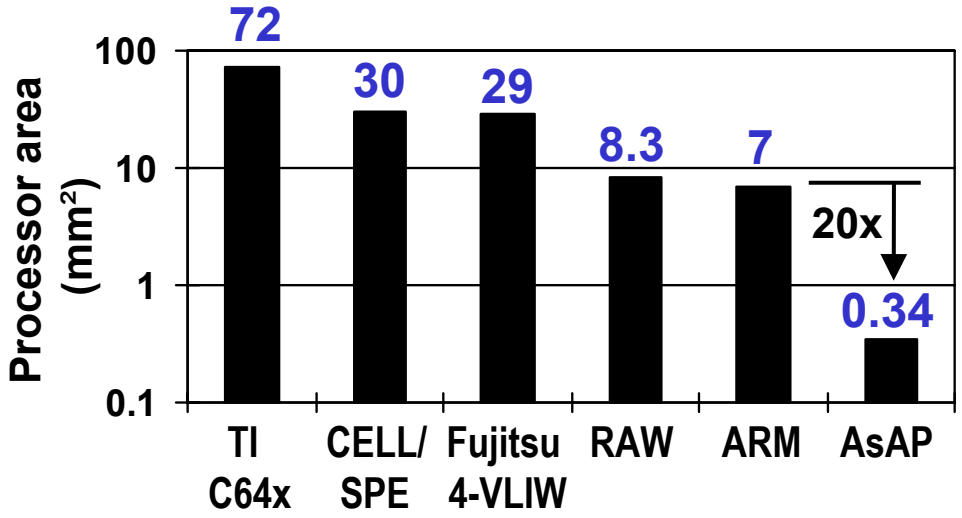
Outline

- Project objectives
- Key features of the AsAP processor
- Design of the AsAP processor
- **Results**

Area Evaluation

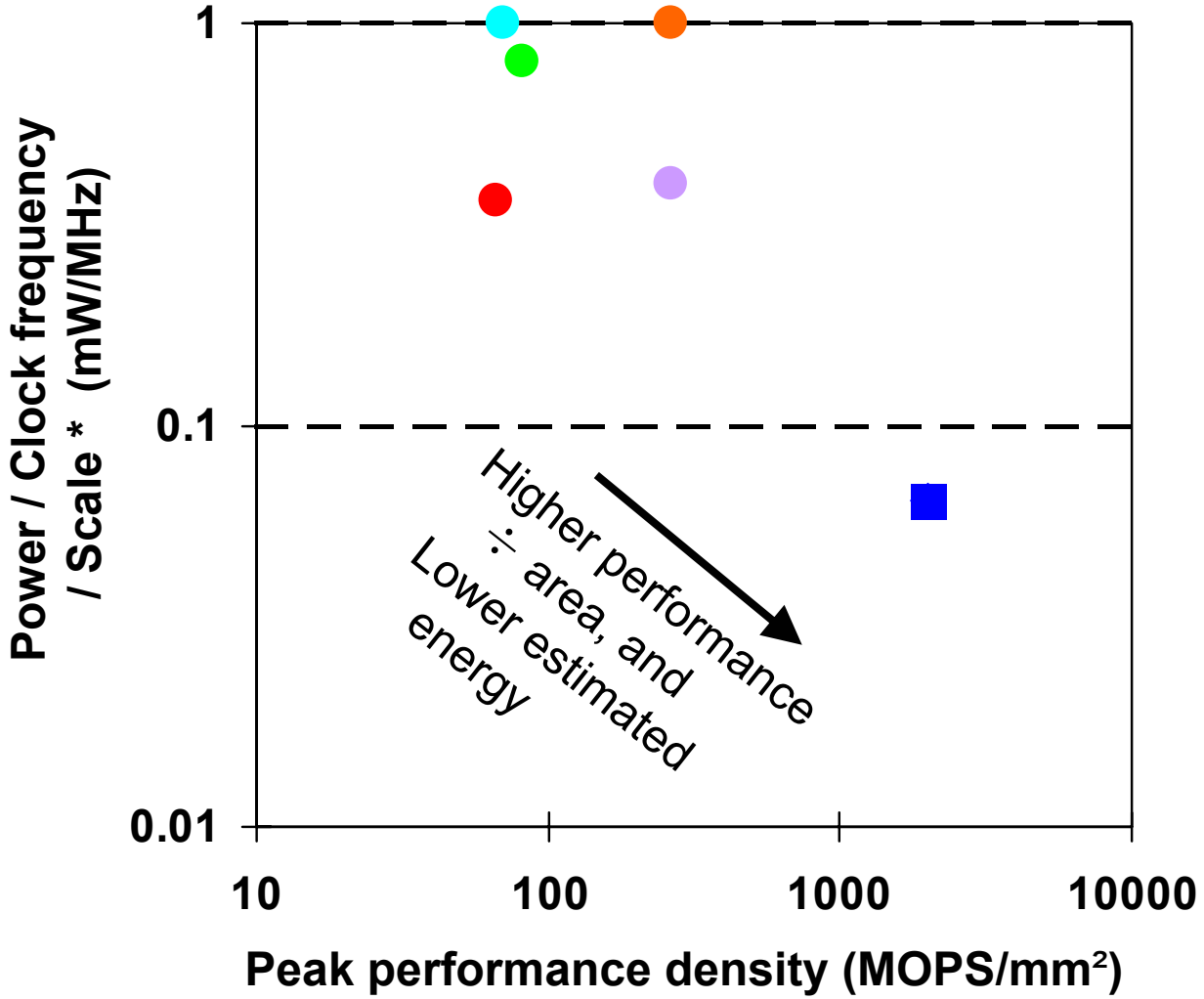
- Most of AsAP's area is for the core (66%)
- Each processor requires a very small area; more than 20x smaller than others





All scaled to 0.13 µm
[ISSCC 05, 00; ISCA04; CMPON96]

Power and Performance



All scaled to 0.13 µm

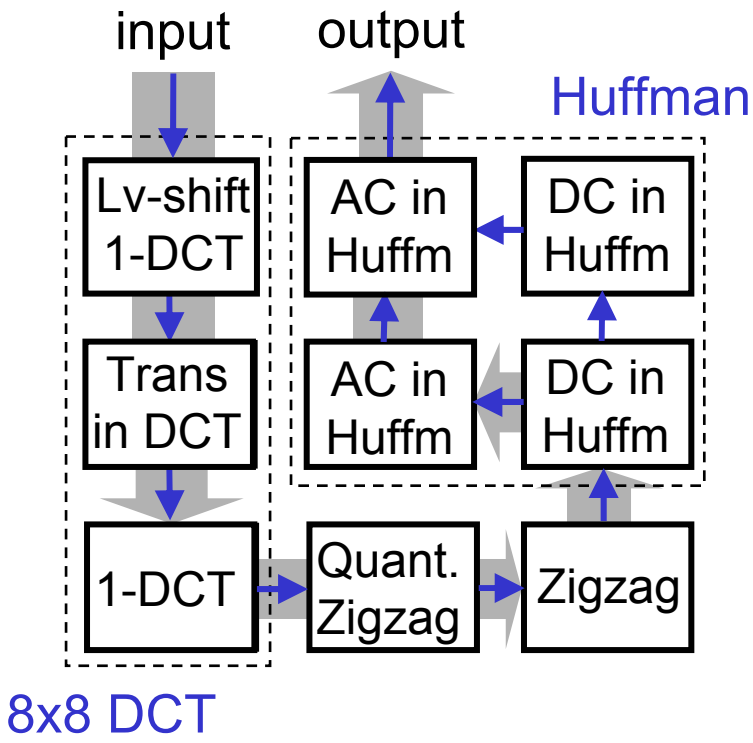
* Assume 2 ops/cycle for CELL/SPE and 3.3 ops/cycle for TI C64x

Note: word widths and workload not factored

[ISSCC 05, 00; ISCA04; CMPON96]

JPEG Core Encoder Implementation

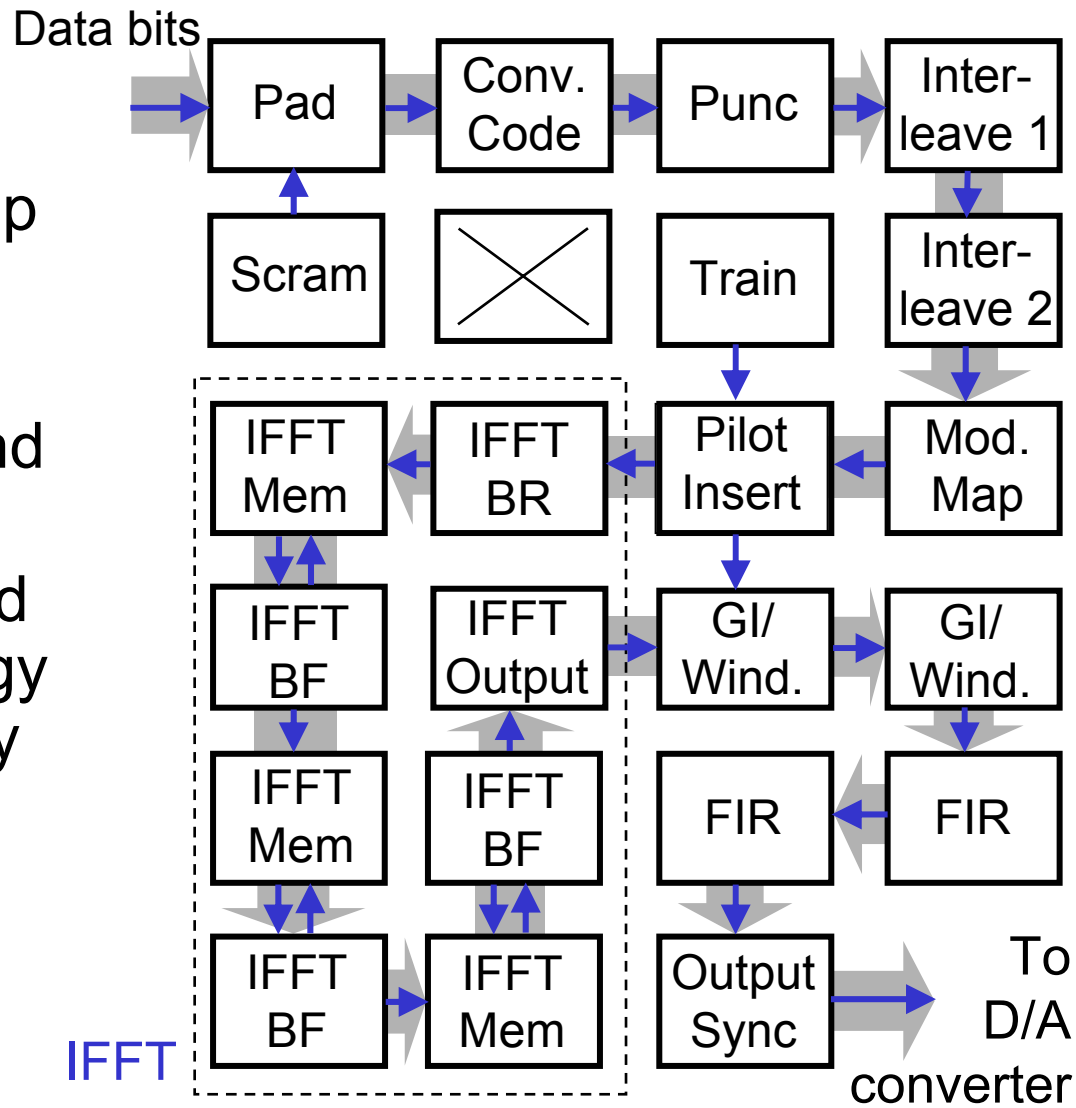
- 9 processors
- Fully functional on chip
- 224 mW @ 300 MHz
- ~1400 clock cycles for each 8x8 block
- Similar performance and ~11x lower energy dissipation than 8-way VLIW TI C62x



[MICRO 02, ISCAS 02]

802.11a/802.11g Wireless Transmitter Implementation

- 22 processors
- Fully functional on chip
- 407 mW @ 300 MHz
  30% of 54 Mb/s
- Code unscheduled and lightly optimized
- ~10x performance and 35x – 75x lower energy dissipation than 8-way VLIW TI C62x



[SIPS 04; ICC 02]

Summary

- Scalable programmable processing array
  - Many processors on a single chip
  - Reduced complexity processors with small memories
  - GALS clocking style
  - Nearest neighbor communication
- Results
  - 0.18 µm, **475 MHz** @ 1.8 V
    - **32 mW** application power
    - **84 mW** 100% active power
    - **2.4 mW** application power @ 116 MHz and 0.9 V
  - High performance density: 475 MOPS in **0.66 mm²**
  - Well suited for future fabrication technologies

Acknowledgments

- Funding
  - Intel Corporation
  - UC Micro
  - NSF Grant No. 0430090
  - MOSIS
  - UCD Faculty Research Grant
- Special Thanks
  - R. Krishnamurthy, M. Anders, S. Mathew, S. Muroor, W. Li, C. Chen, D. Truong