



SmartCell Reconfigurable Architecture for Low-Power Stream Processing

Cao Liang and *Xinming Huang*
Embedded Computing Lab
Worcester Polytechnic Institute
<http://computing.wpi.edu>

MAPLD Conference
September 15-18, 2008
Annapolis, MD



Outline

- Introduction and Motivation
- SmartCell Architecture
- SmartCell Prototype with 64 PEs
- Benchmark Applications and Performance
- Conclusions

Challenges

- Major driving force in embedded computing
 - Multimedia signal and image processing
 - Wireless communications
 - Military and space applications
- Design challenges:
 - Low power (power efficiency)
 - High performance
 - Flexibility (Programmability or reconfigurability)



Multimedia TV



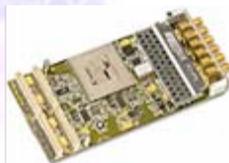
Game Console



Data encryption



Scientific computing



Software radio



PDA



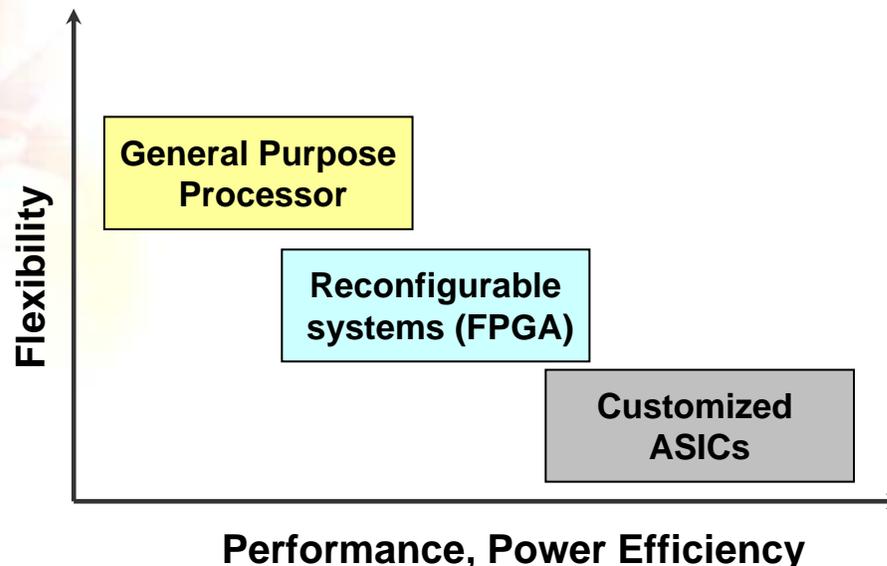
Image processing



Radar imaging

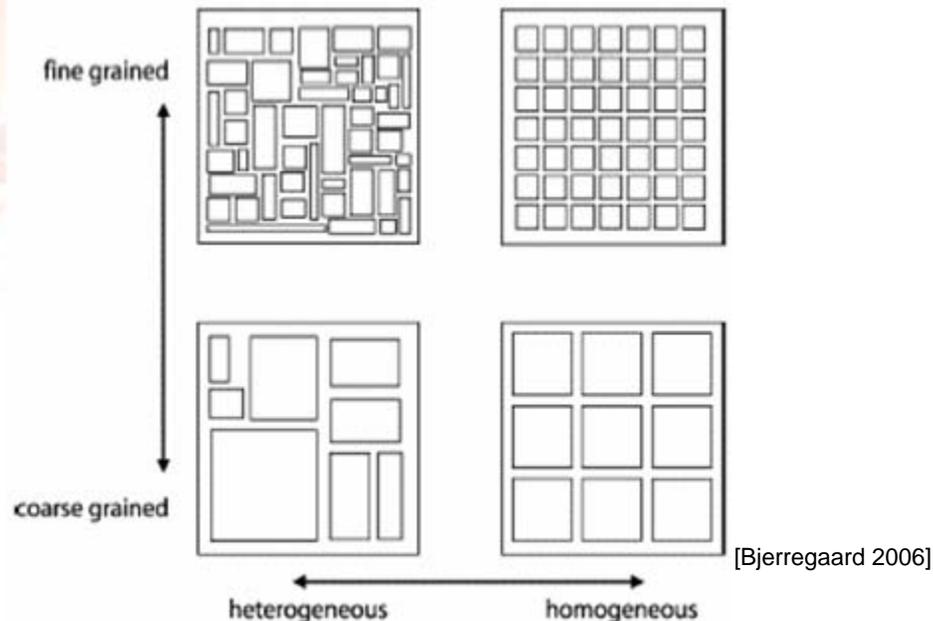
Existing Computing Platforms

- General purpose processors (GPP)
- Application specific integrated circuit (ASIC)
- Reconfigurable architecture
 - Dominated by Field Programmable Gate Array (FPGA)
- New architectures: CellBE, GPU



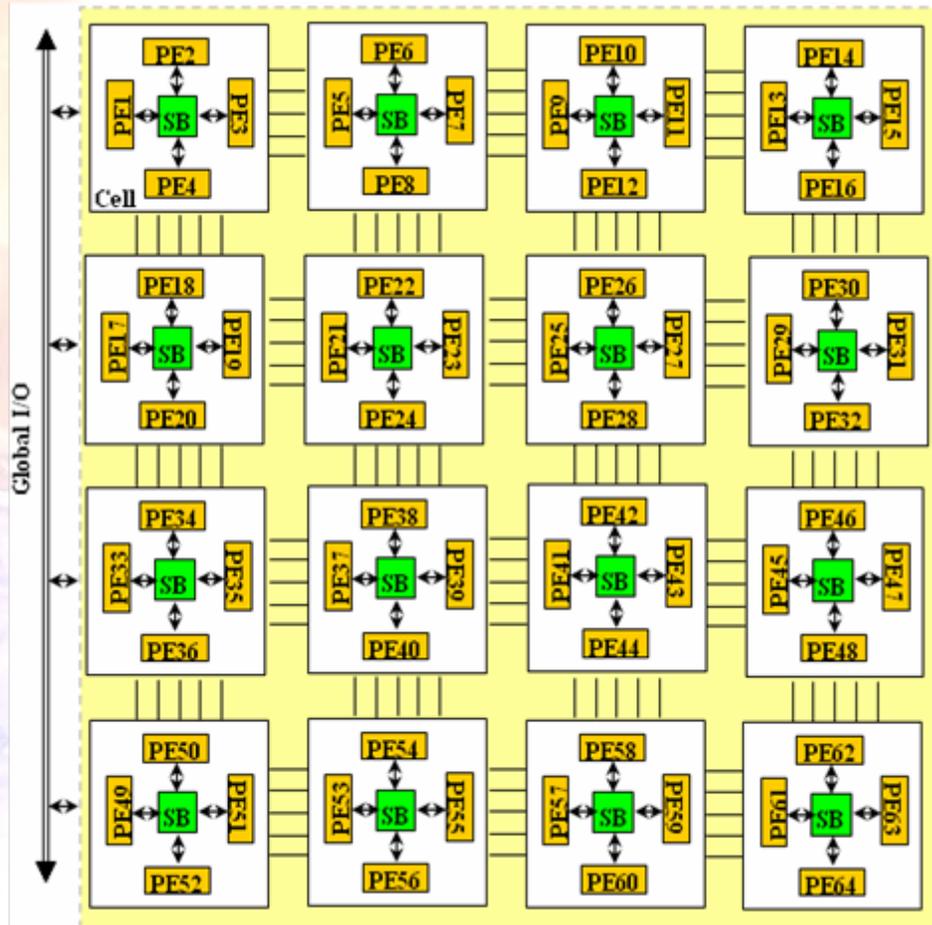
Coarse-Grained Reconfigurable Architecture (CGRA)

- Motivations of the SmartCell architecture
 - Coarse-grained computing operators
 - Reconfigurable interconnection
 - Domain specific, e.g. stream processing
- Bridging the gap between FPGA and ASIC



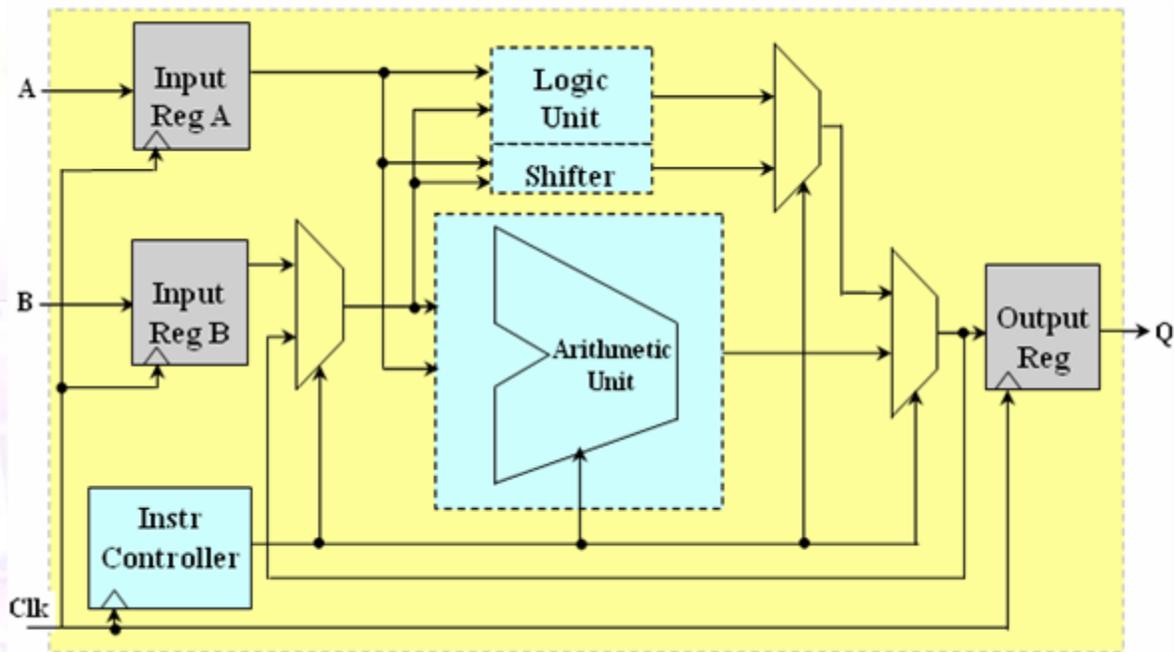
Overview of SmartCell Architecture

- Computing units are tiled in a 2D structure



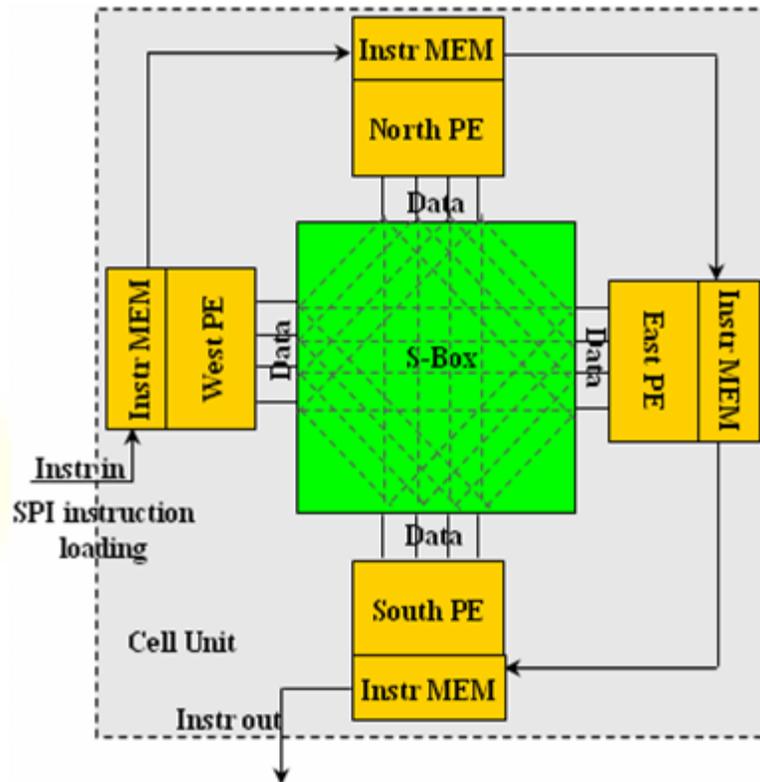
Design of Processor Element

- Processor Element (PE)
 - 16-bit input, 36-bit output
 - Logic, Shift, and Arithmetic operations



Design of Cell Unit

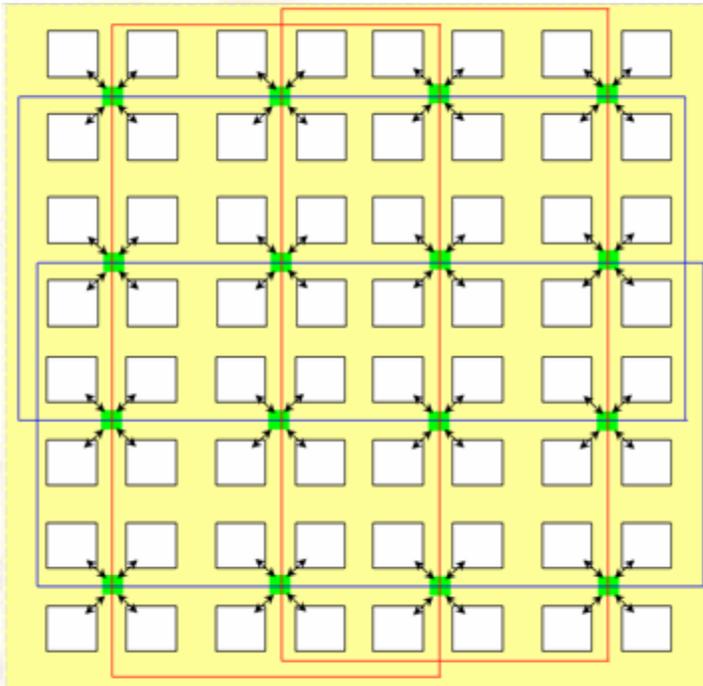
- Include 4 PEs to form a **quad** structure
- Fully connected cross-bar (S_Box) for data exchange
- Serial peripheral interface (SPI) for instruction configuration



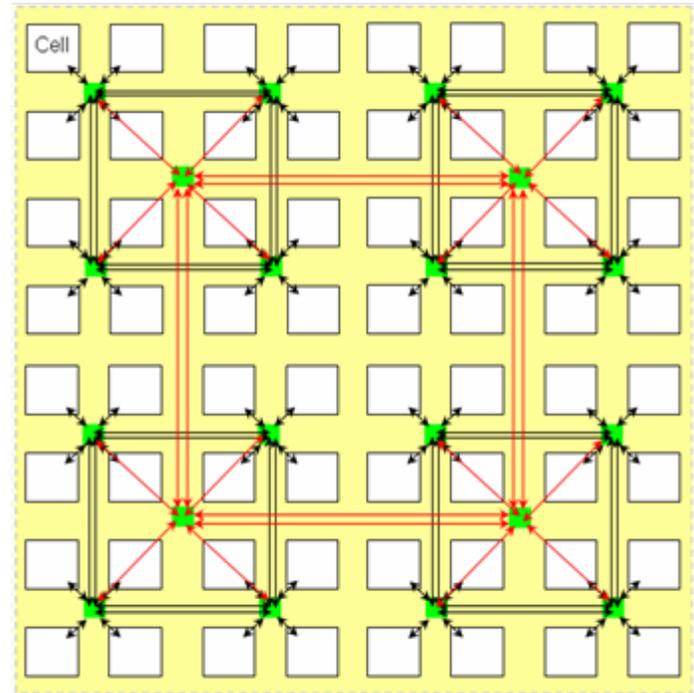
On-chip Interconnection Design

- Modified CMesh On-chip Network

Traditional CMesh



Hierarchical CMesh



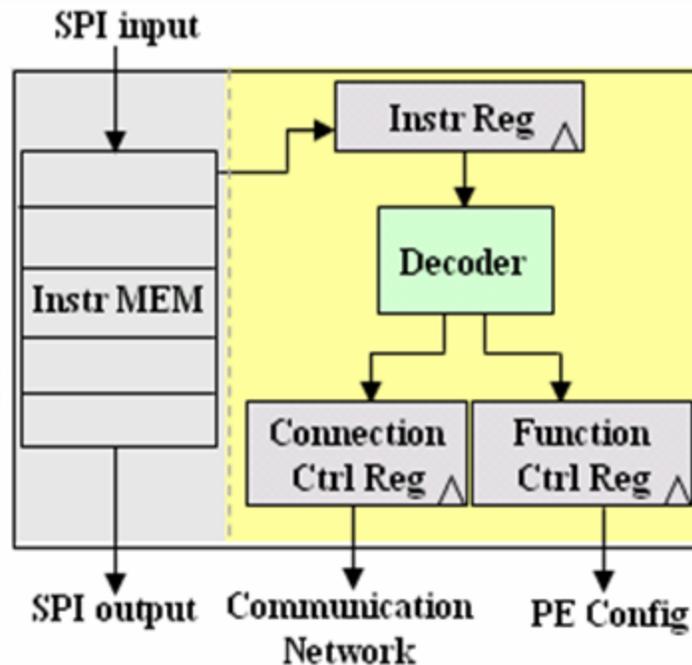
Control Logic Design

- Four types of control signals
 - Program counter control
 - Datapath/delay control
 - Operation control
 - Network-on-Chip control
- Format of the instruction code (64-bit/instruct)

64 bits/instruction code						
# of bits	9	20	7	10	11	7
Format	PC control	Datapath control	I/O delay	Operation control	NoC control	RESV

Configuration Structure

- System configuration

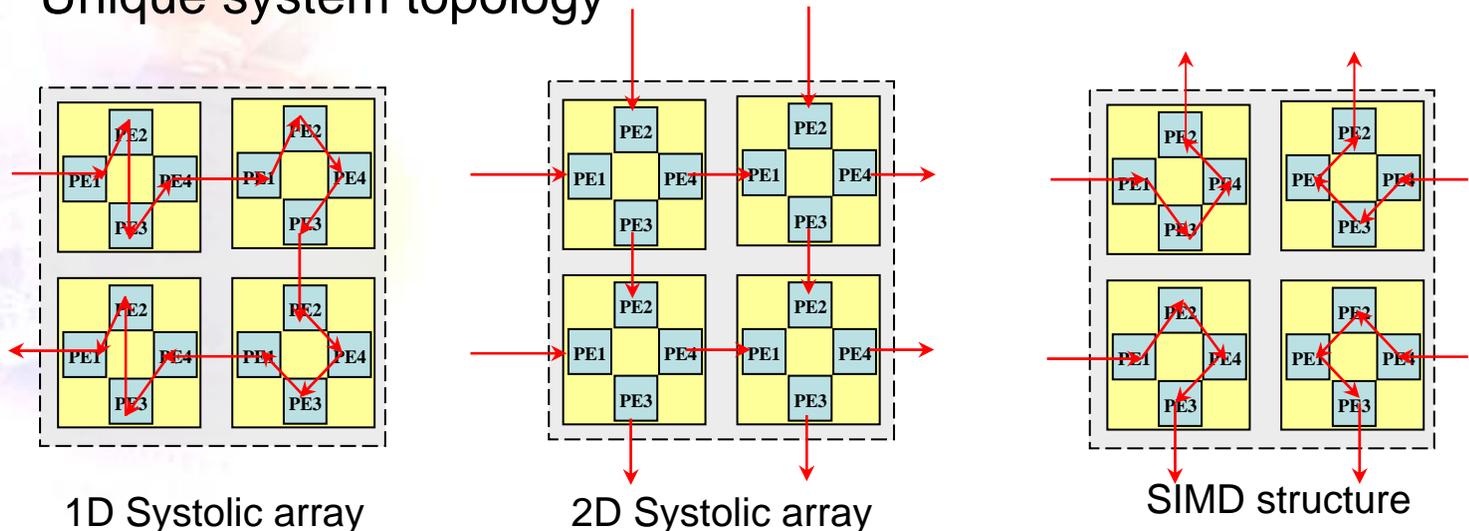


Prototype Chip Design

- Implementation of a seedling SmartCell system with 16 cell units in a 4 by 4 mesh structure, with a total of 64 PEs
 - RTL level design and simulation
 - FPGA prototyping
 - Standard cell ASIC implementation with TSMC .13 μ m technology
 - Total area is about 8.2 mm²
 - Runs up to 107 MHz
 - Configuration time is within 12 μ s

SmartCell Features

- A combination of the following features makes SmartCell a unique approach in CGRA families
 - Dynamic reconfiguration
 - Deep pipeline and parallelism
 - Hardware virtualization
 - Explicit synchronization
 - Unique system topology



Application Domain and Benchmarks

Application Domain	Test Benches
Signal processing	64-tap FIR 64-tap IIR
Multimedia and image processing	32-point FFT 8*8 2D-DCT, 8 by 8 Motion Estimation (ME) in 24 by 24 searching area
Scientific computing	128 by 128 Matrix Multiplication (MMM), 64 th -order Polynomial Evaluation (PoE) RC5 Data Encryption

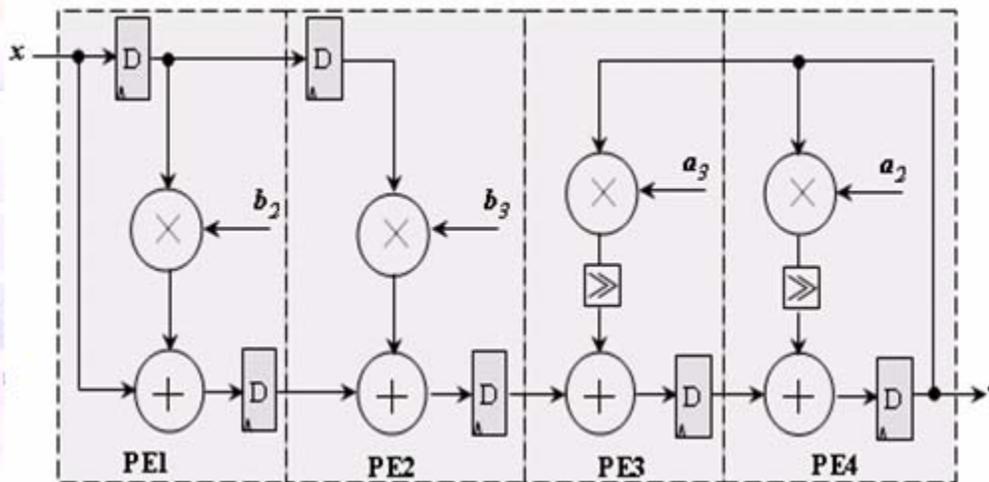
Benchmark Mapping

- Infinite Impulse Response (IIR) filter

$$y(n) = \sum_{i=1}^N a_i y(n-i) + \sum_{i=0}^M b_i x(n-i)$$

- Biquad cascaded-IIR structure on a single Cell

$$H(z) = \prod_{i=1}^K H_i(z), \text{ where } H_i(z) = \frac{1 + b_{i2}z^{-1} + b_{i3}z^{-2}}{1 + a_{i2}z^{-1} + a_{i3}z^{-2}}$$

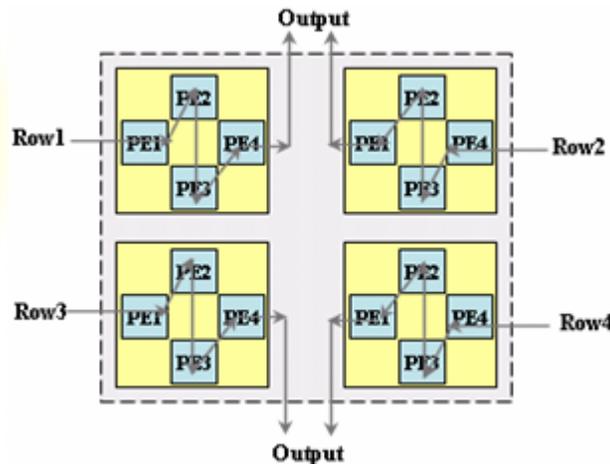
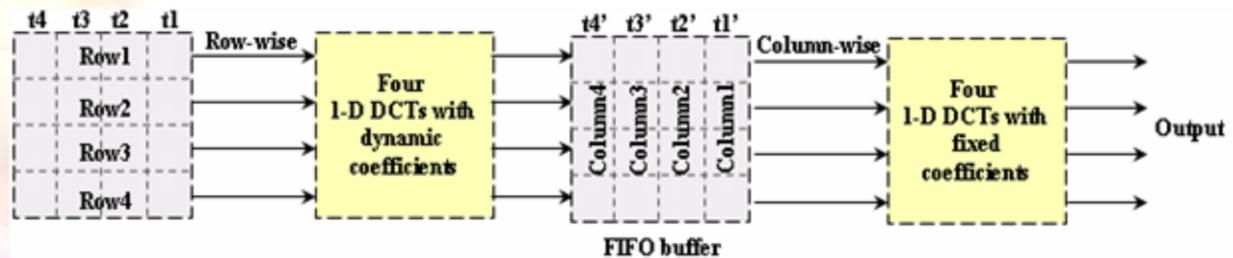


Benchmark Mapping (cont')

- 2D Discrete Cosine Transform (2D DCT)

$$X_{i,j} = a_i b_j \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} x_{k,l} \cos \left[\frac{\pi}{N} (k+1/2)i \right] \cos \left[\frac{\pi}{N} (l+1/2)j \right], \text{ where } 0 \leq i, j < N$$

- Decomposed into two 1D DCTs



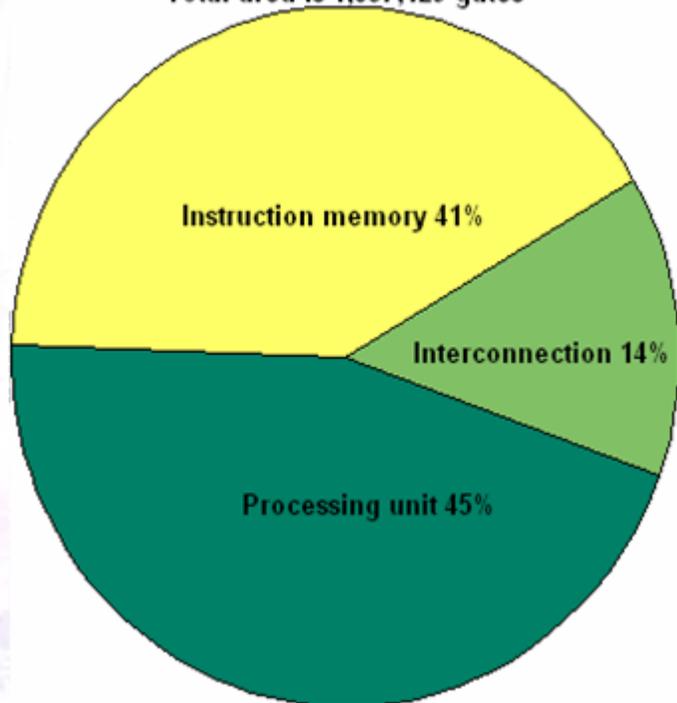
Experimental Setup

- Evaluation Metrics
 - Area & Timing
 - Power consumption
 - Throughput and power efficiency
 - Comparing with RaPiD, Altera's Stratix II FPGA and ASIC

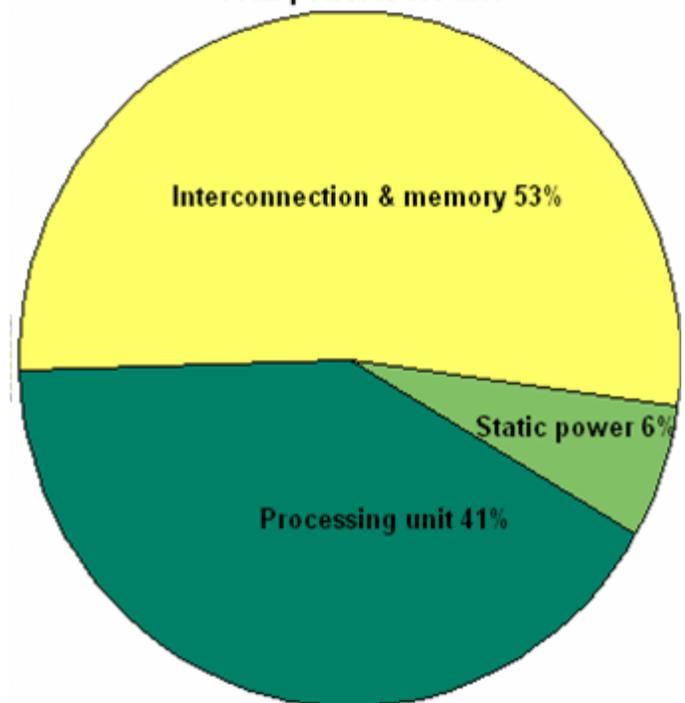
System dimension	4 by 4
Design tools	ModelSim, Synopsys
Library	TSMC .13 μ m process
Voltage	1 V
Simulation freq.	100 MHz

Area and Power Consumption

Total area is 1,637,129 gates



Total power is 156 mW



Generated at 100 MHz with fully operational circuits

Power Consumption and Efficiency

- On average 156 mW power consumption @ 100 MHz
- 31 GOPS/W energy efficiency
 - only arithmetic & logic operations, excluding I/O power

	FIR	IIR	2D-DCT	RC5	MMM	FFT	PoE	ME
P_{Dyn} <mW>	144	181	153	132	135	161	142	137
P_{Core} <mW>	152	189	161	140	143	169	150	145
E_{Eff} <GOPS/W>	42.1	33.9	39.8	45.7	11.2	18.9	42.7	11.0

Compare with RaPid, FPGA, and ASIC

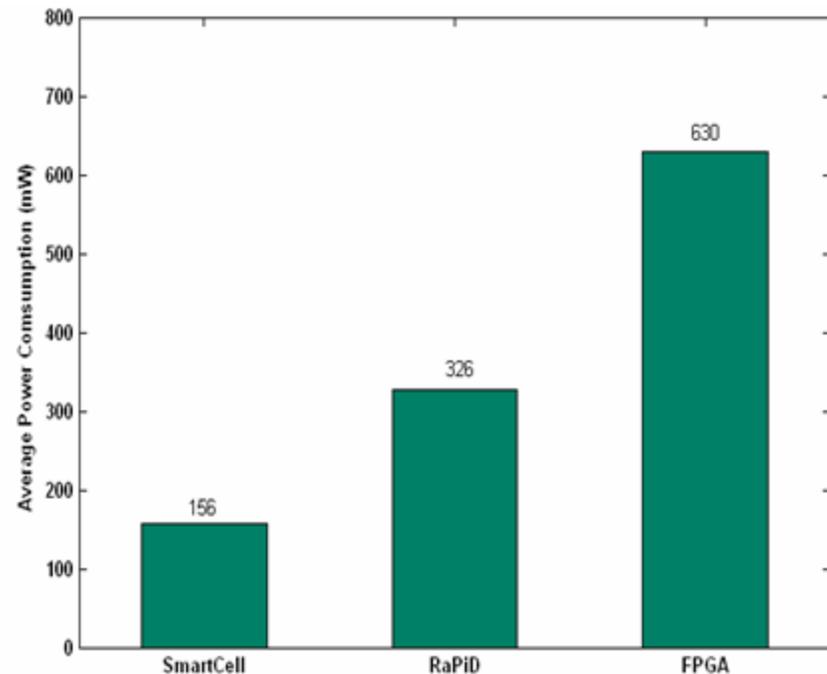
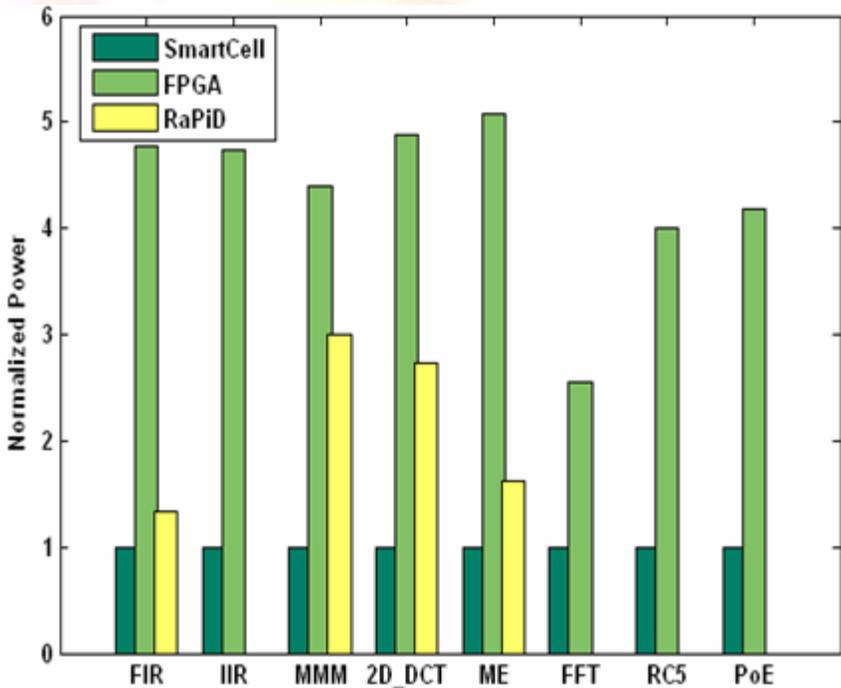
- **Power and system throughput comparison**

- Power consumption of RaPiD has been scaled down to the same process technology of SmartCell system

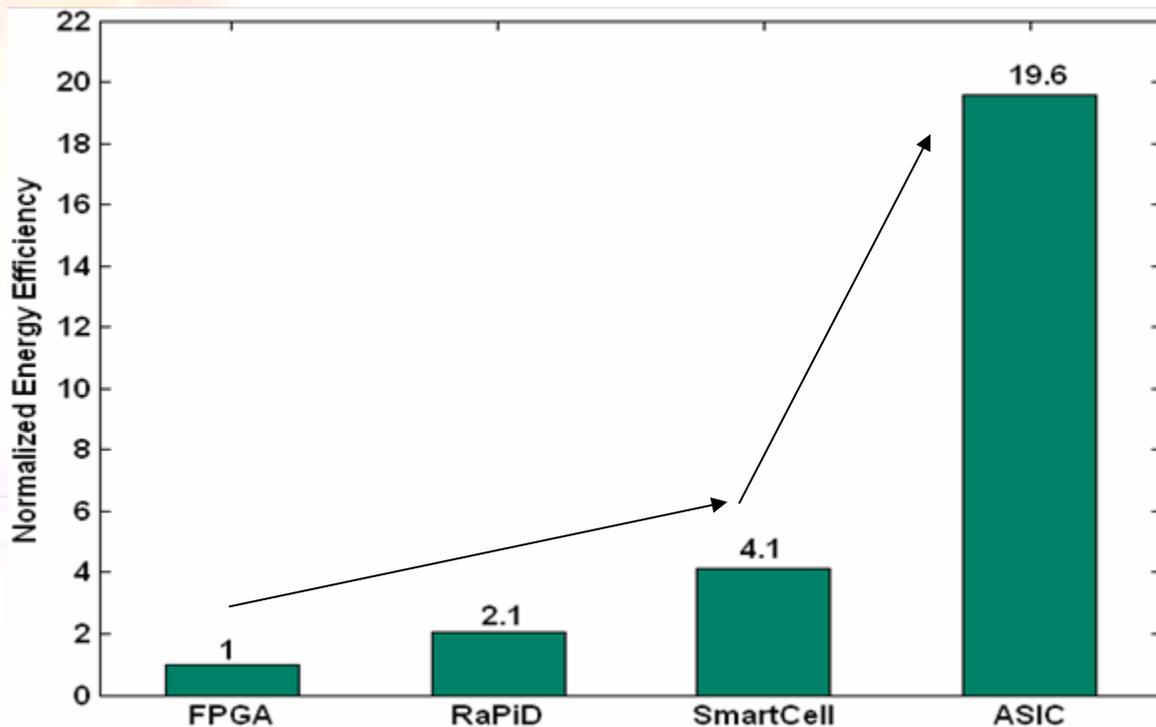
		FIR	IIR	MMM	2D DCT	ME	FFT	PoE
SmartCell	Power <mW>	152	189	143	161	145	169	150
	Throu.*	100 MS/s	100 MS/s	763 Matrices/s	1.56 Mblocks/s	865 Kblocks/s	58 MS/s	100 MS/s
RaPiD [26]	Power <mW>	203	-	428	439	235	-	-
	Throu.	100 MS/s	-	763 Matrices/s	1.56 Mblocks/s	865 Kblocks/s	-	-
FPGA	Power <mW>	725	896	445	787	573	431	628
	Throu.	100 MS/s	100 MS/s	763 Matrices/s	1.56 Mblocks/s	865 Kblocks/s	100 MS/s	100 MS/s
ASIC	Power <mW>	31	45	9	55	12	33	55
	Throu.	100 MS/s	100 MS/s	763 Matrices/s	1.56 Mblocks/s	865 Kblocks/s	58 MS/s	100 MS/s

Compare with Rapid and FPGA

- 52% average power reduction compared with RaPiD
- 75% average power reduction compared with FPGA



Power Efficiency Comparison



* Compare to 90nm Stratix-II FPGAs

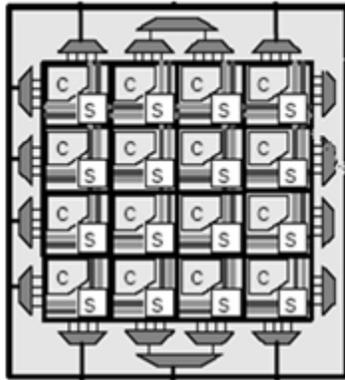
Conclusions

- An interesting CGRA architecture is proposed and developed – namely SmartCell
- The architecture is reconfigurable and can be targeted for different computing systems
- A prototype design with 64 PEs shows both throughput and power efficiency in benchmarks of data streaming applications
- SmartCell may have the potential to bridge the gap between high-power FPGAs and inflexible ASICs

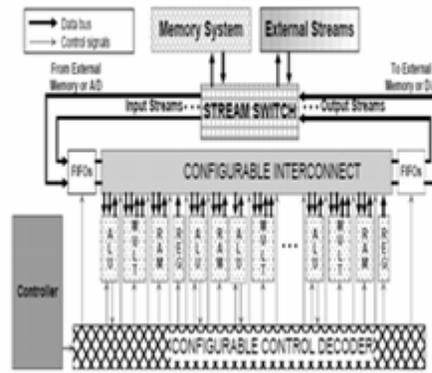
The following are backup slides



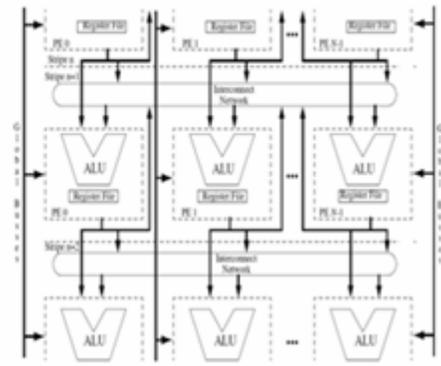
Existing CGRAs



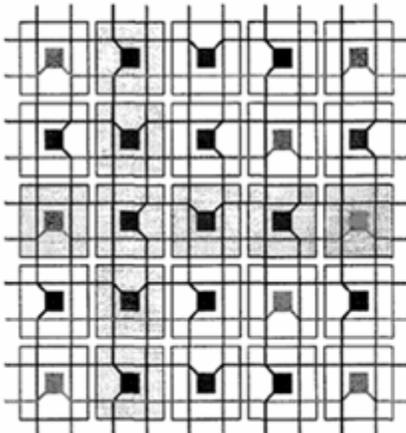
RAW 2D mesh structure [MIT]



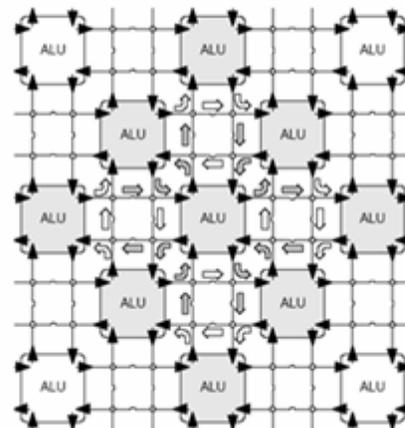
RaPiD 1D array [UW]



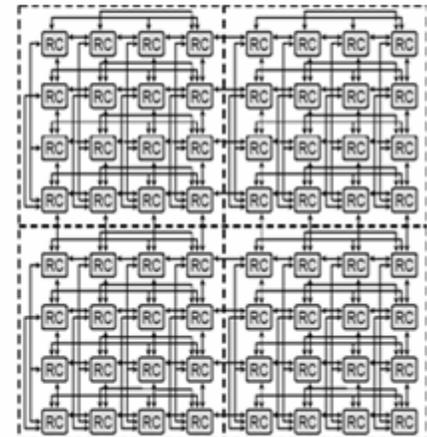
PipeRench 1D Stripe architecture [CMU]



Matrix 2D mesh structure [MIT]



Chess hexagon structure [HP Lab]



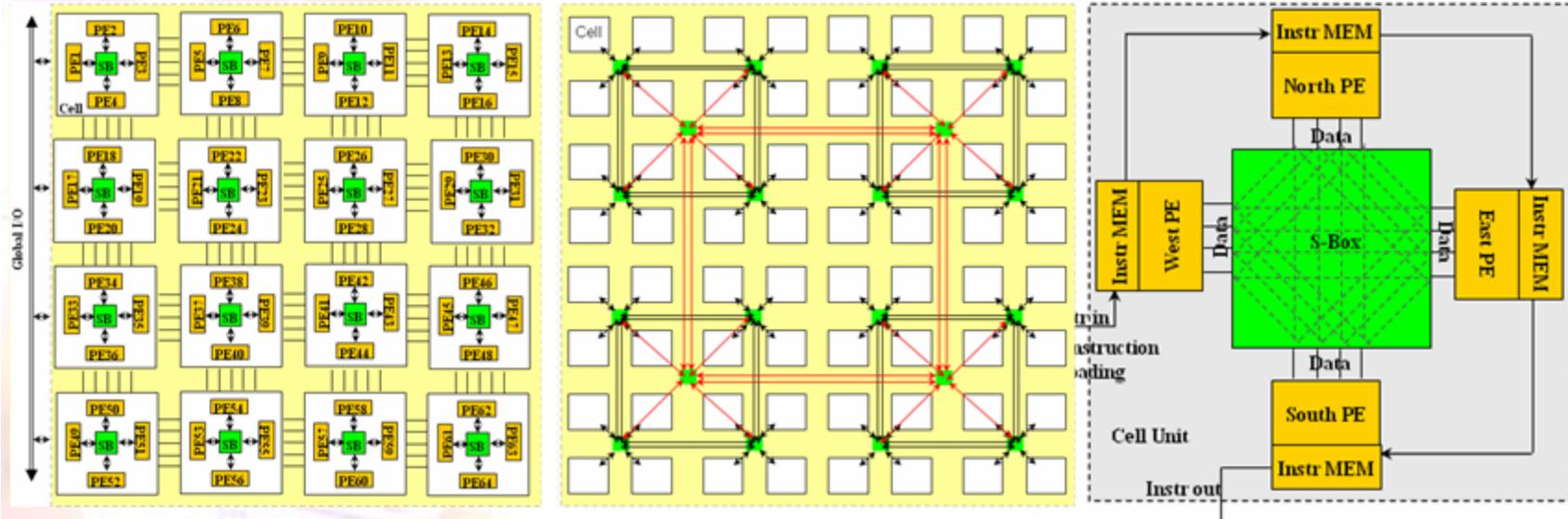
MorphoSys mesh structure [UC_Irvine]

Characteristic Comparison

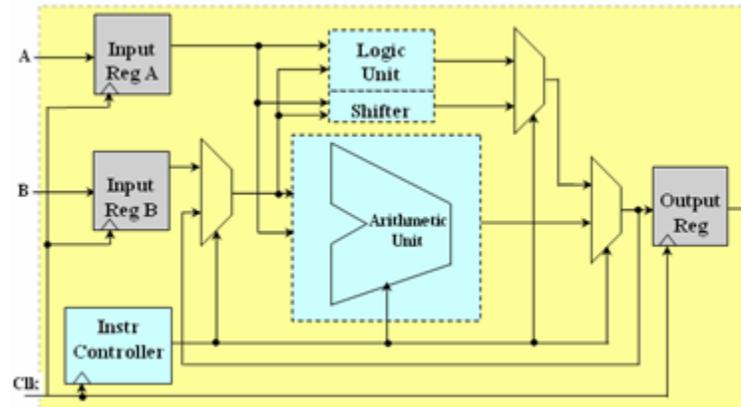
Characteristics of the compared CGRAs

System	Archit.	Reconfig.	Application	Connection	Homogeneity	Parallel & Pipeline
RAW	MIMD	Static	Irregular or general	2D Mesh	Homogeneous	Spatial & Temporal
RaPiD	SIMD	Mixed	Systolic or pipeline	1D array	Heterogeneous	Temporal
PipeRench		Dynamic	Stream-based	1D array	Homogeneous	Spatial & Temporal
MATRIX	VLIW SIMD MIMD	Dynamic	Systolic	Layered Structure	Homogeneous	-
CHESS		Static	Multimedia	Hexagon mesh	Homogeneous	-
MorphoSys	SIMD	Dynamic	Data-parallel	2D mesh	Homogeneous	Spatial
SmartCell	SIMD MIMD Systolic	Mixed	Multimedia or Systolic	Layered Structure	Homogeneous	Spatial & Temporal

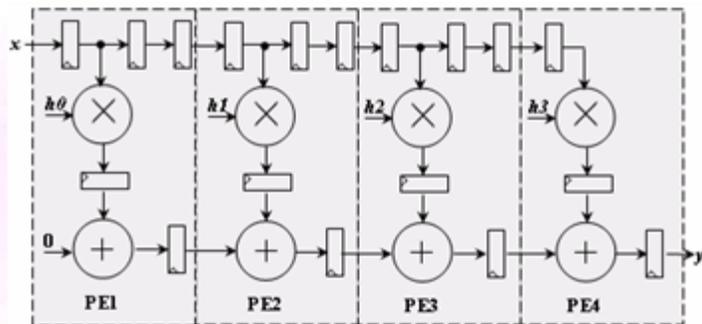
SmartCell: Tiled Architecture, Processor Design and Interconnect



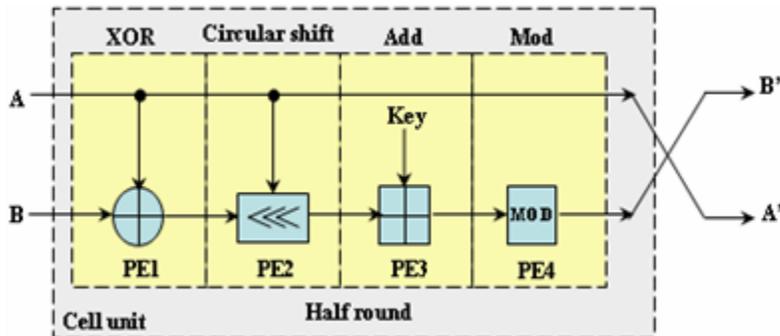
- Many cells are tiled in 2D layout
- Each cell has 4 PEs (N,W,S,E)
- Simplified processor with mem
- A crossbar within the cell; on-chip interconnect uses CMesh



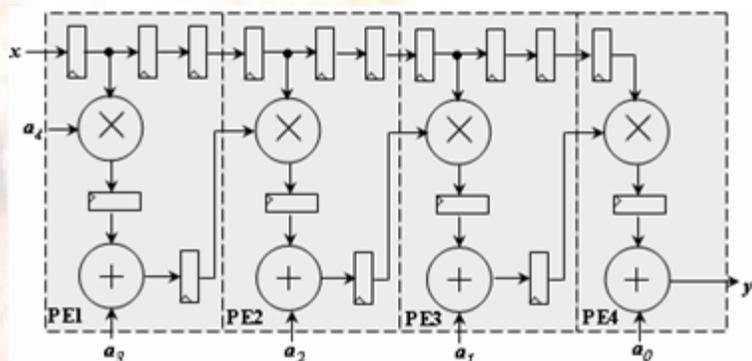
Features and Application Domain



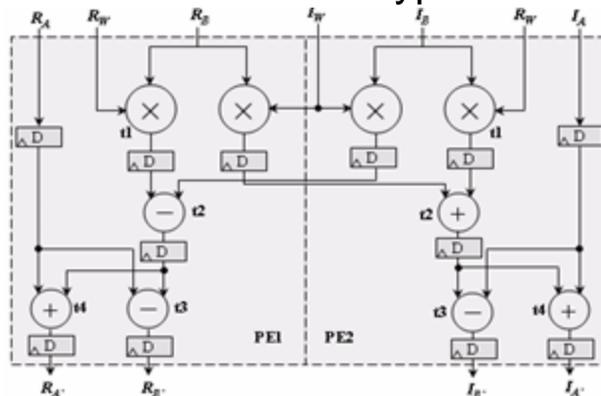
FIR filter



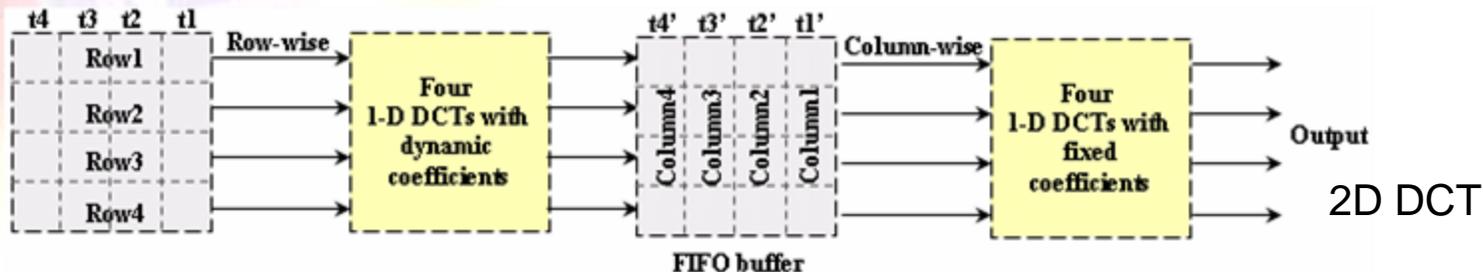
RC5 data encryption



Polynomial Evaluation



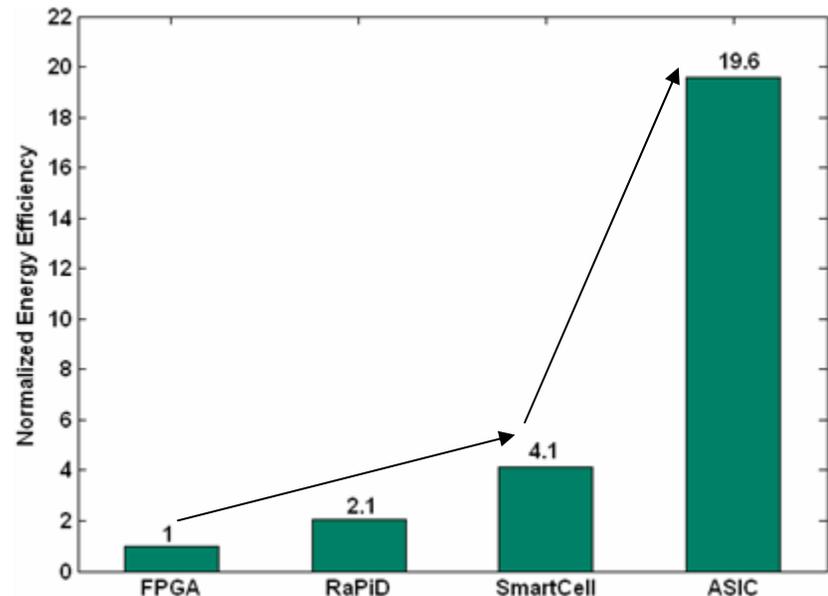
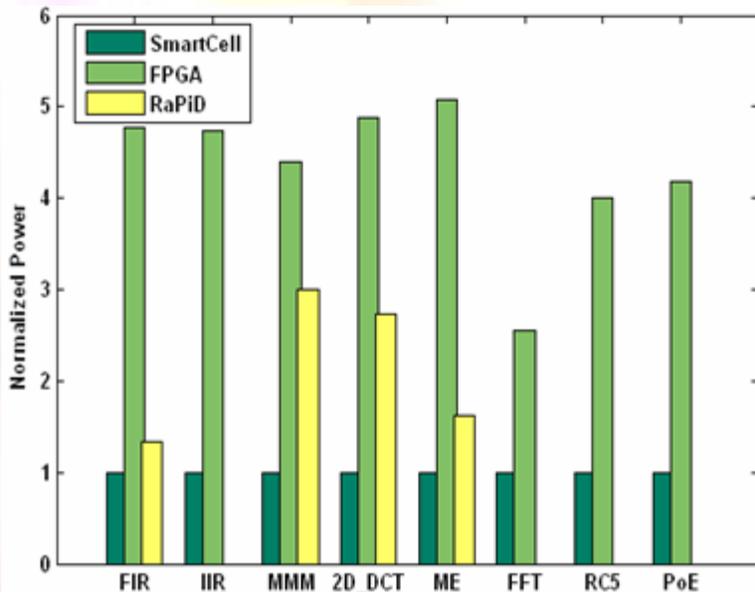
FFT Butterfly



2D DCT

System Design and Performance

- Prototype chip design: 4x4 cells (64 PEs), .13 TSMC, 8.2mm², 1V, about 156mW @100MHz
- Benchmark with RaPiD, Stratix-II (90nm), and ASIC



Acknowledgement:

- Dr. Michael Fritz, DARPA/MTO YFA Program
- Cao Liang, WPI graduate assistant, now with AMD