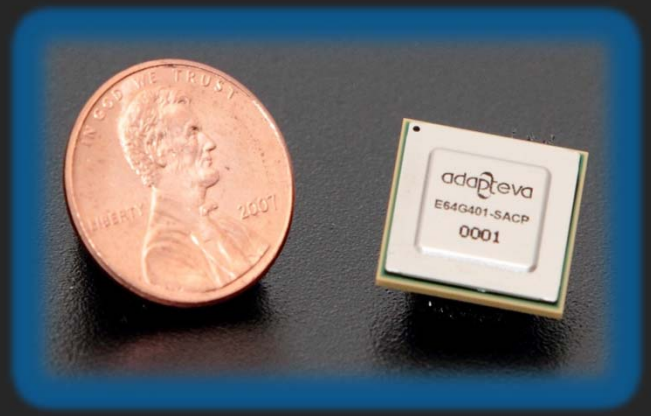
A Closer Look at the Epiphany-IV 28nm 64-core Coprocessor



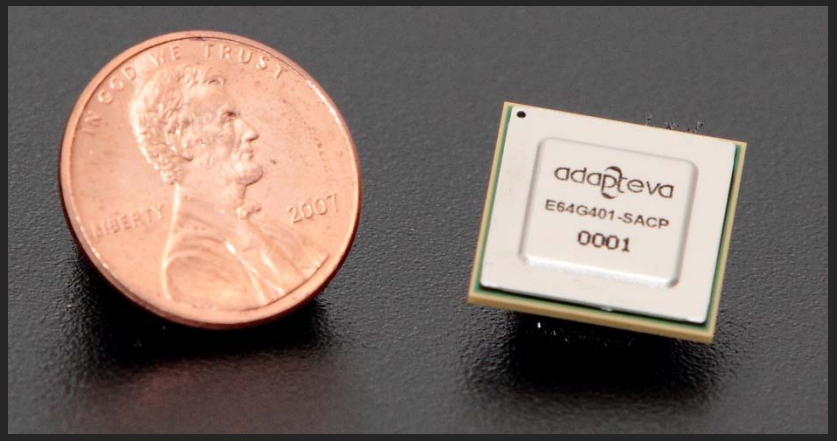
Andreas Olofsson

PEGPUM 2013

adapteva

Adapteva Achieves 3 “World Firsts”

1. First processor company to reach 50 GFLOPS/W



2. First mobile processor with an open source OpenCL™ SDK

3. First semiconductor company to successfully crowd-source project

KICKSTARTER

adapteva

The Future is…

Efficient…

Heterogeneous…

Open..

Task Parallel…

Grande Challenges Ahead…

Rebuild the computer ecosystem!

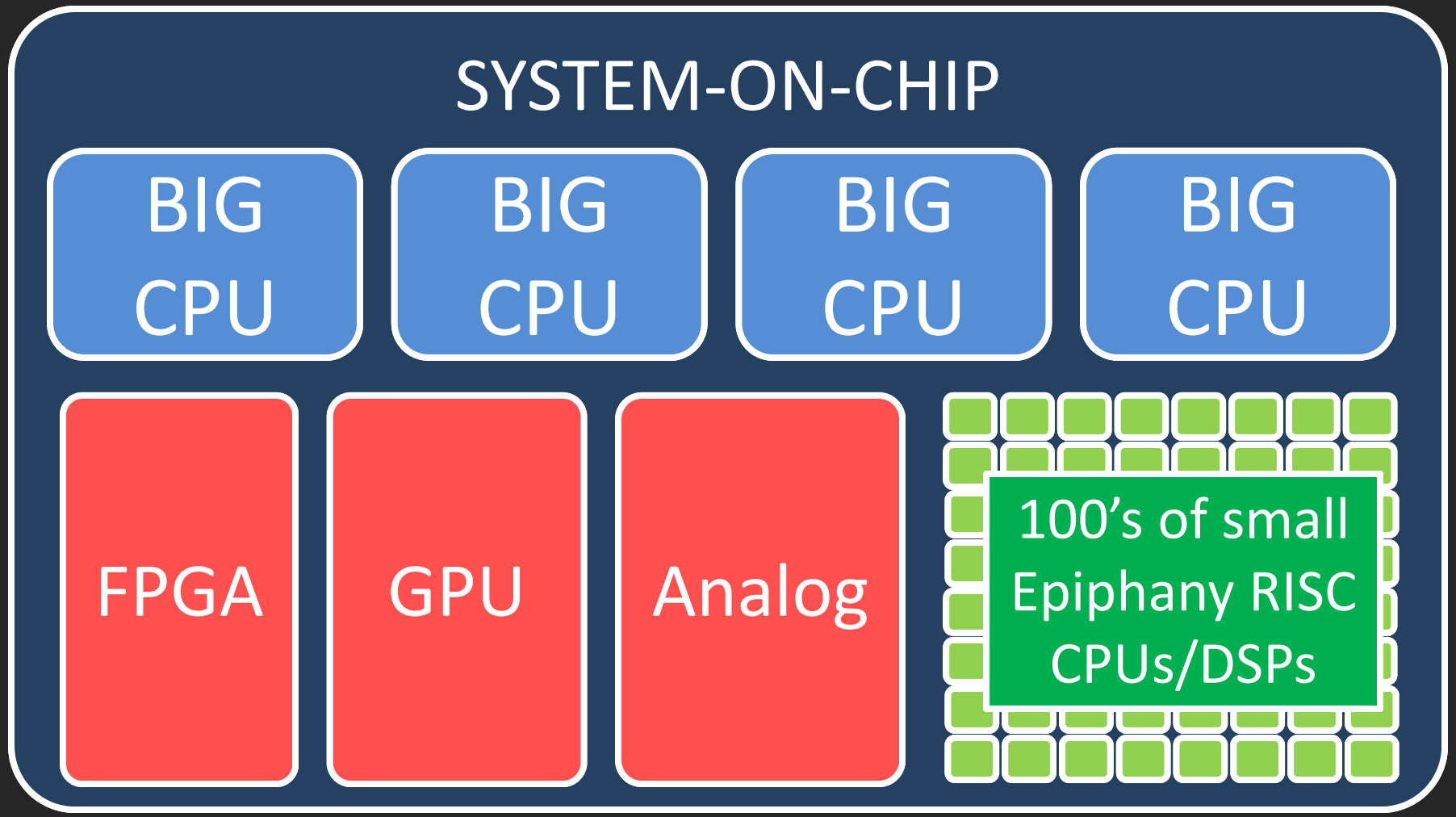
Rewrite billions of lines of code!

Retrain millions of programmers!

Rewrite the education curriculum!

adapteva

Our Vision: True Heterogeneous Computing



adapteva

Architecture Comparison

Technology	FPGA	DSP	GPU	CPU	Manycore
Process	28nm	40nm	28nm	32nm	28nm
Programming	VHDL	OCL/C++/C	CUDA/OCL	OCL/C/C++	OCL/C/C++
Area (mm^2)	**590**	**108**	**294**	**216**	**10**
Chip Power (W)	40	22	135	130	2
CPUs	n/a	8	16	4	64
Max GFLOPS	**1500**	**160**	**3000**	**115**	**102**
GHz * Cores	n/a	12	16	14.4	51.2
Compile Time	Hours	Minutes	Minutes	Minutes	Minutes
L1 Memory	**6MB**	**512KB**	**2.5MB**	**256KB**	**2MB**

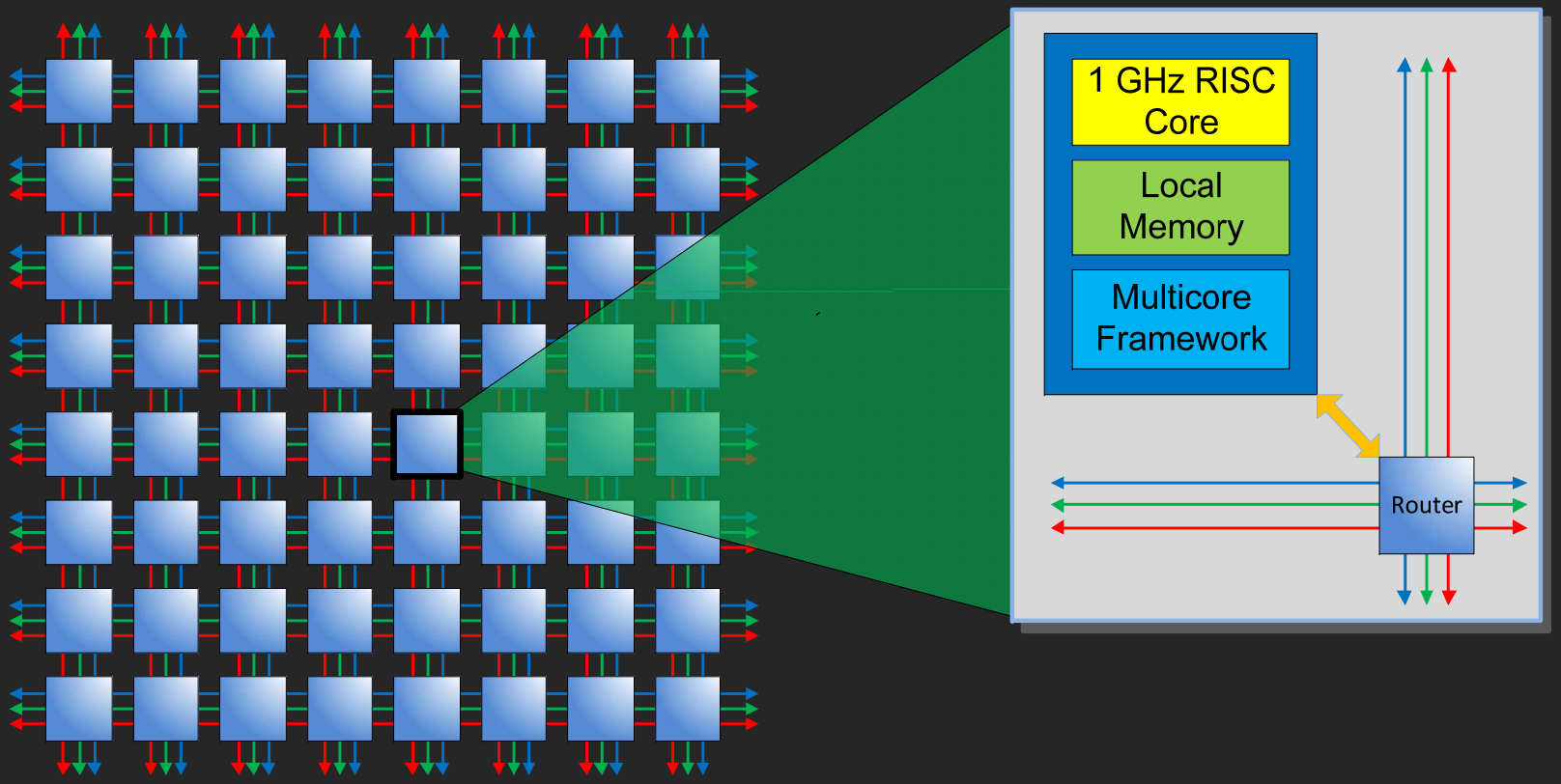
Efficiency is everything

Peak performance means very little

No magic bullet!

adapteva

Epiphany: Massive Task-Parallelism



Coprocessor to ARM/Intel CPU

25mW per core

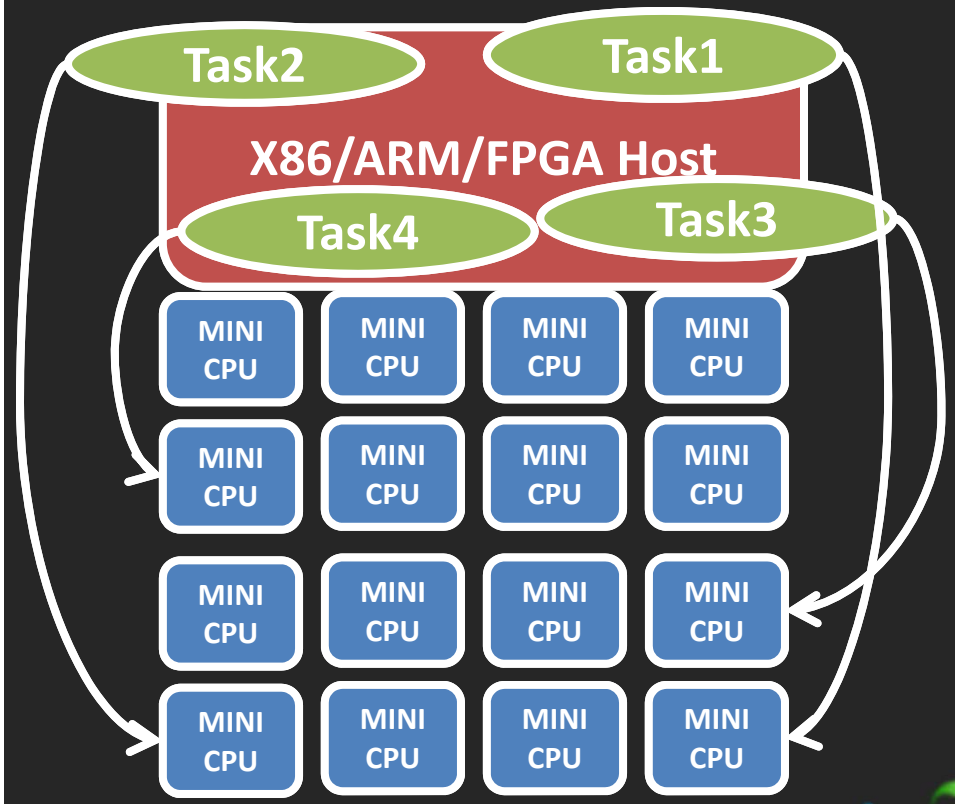
C/C++ programmable

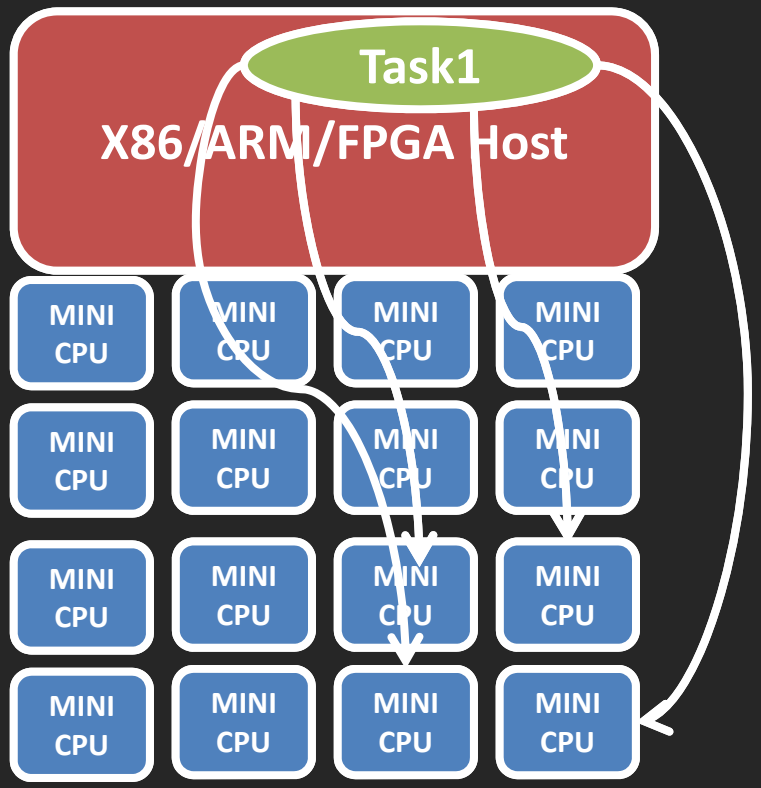
adapteva

Programming Models

MODEL#1
TASK QUEUE MODEL
- Great for up to 2GFLOPS
- Supports standard C/C++
- "Cloud on a chip"

MODEL #2
DATA PARALLEL MODEL
- openCL programmable
- Easy integration of C/C++
- openMP/MPI roadmap





adapteva

CPU Architecture Tradeoffs

Feature	In	Out	Why
Single Precision Floating Point	X		Programming Efficiency
64 Entry Register File	X		GCC
Byte Addressable	X		GCC
64bit load/store	X		Data movement is king
Dual Issue In Order Scheduling	X		
4-Bank Local Memory	X		Multicore data movement
Interrupts, breakpoints, timers	X		Inexpensive
Hardware Cache		X	Too expensive, software managed
Double Precision Floating Point		X	Overkill for initial market
VLIW		X	GCC, Complexity
Instr: Not, Mask,Rotate, Add-Carry, Ones		X	Not important
Data Type ISA Orthogonally. (Signed ,unsigned) & (8b,16b,32b)		X	Expensive. Focus was on floating point

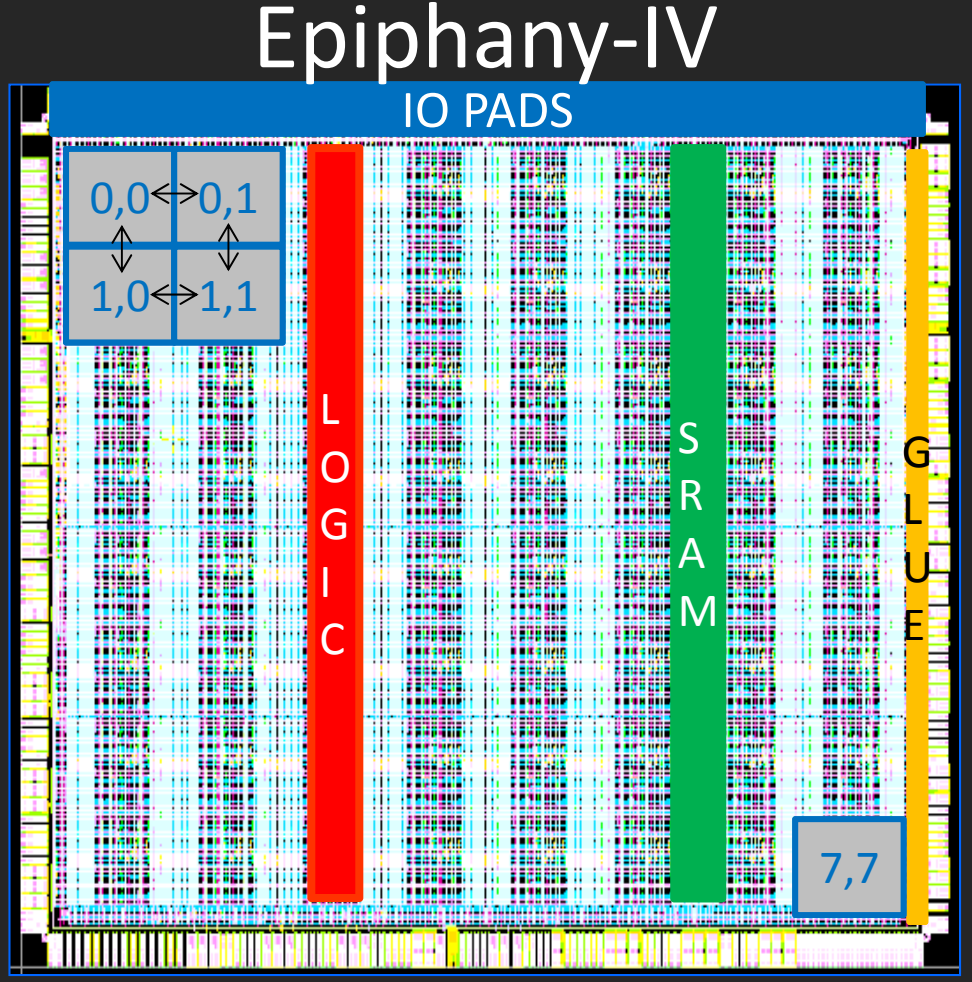
Network-On-Chip Tradeoffs

Feature	In	Out	Why
64-Bit Transfer /Cycle	X		Maximize efficiency
Send address on every cycle	X		“Wires are free”, simplicity
Round robin arbitration	X		Easy, good enough
Distributed routing	X		Scalable
Robust flow control	X		Must have
Bidirectional Mesh	X		Well known
Address mapped packet switching	X		Ease of Use
Single cycle message transfer	X		Short messaging important
3 separate Networks	X		No deadlock, QOS
Extensive QOS features		X	Deterministic traffic
Torus wraparound connections		X	Too expensive in CMOS
Circuit switching		X	Not general purpose
Large routing buffers		X	Too expensive
Multilayer Mesh		X	Too expensive (for now)

adapteva

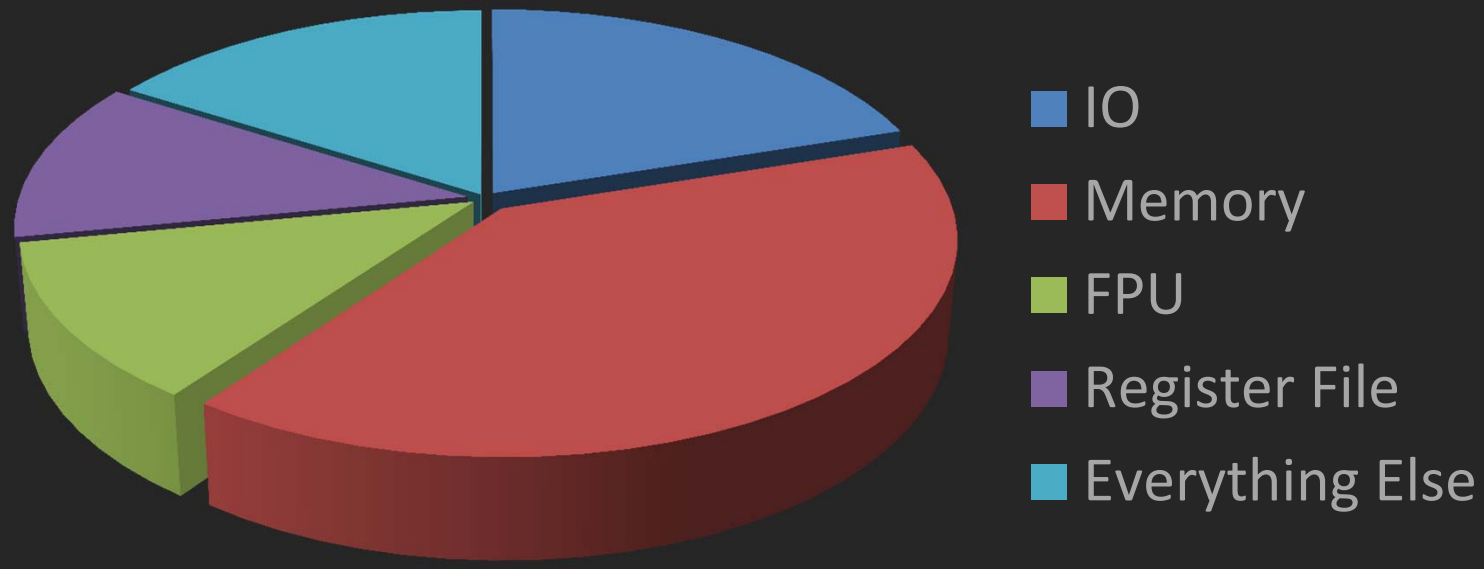
Epiphany Implementation Methodology

- Scalable
- 24 hrs from RTL to GDS
- Reusable tiles
- No long wires
- Pattern density
- Fault tolerant
- Thermal balancing
- Easy clocking



adapteva

Epiphany-IV Area Breakdown



Memory/RF/FPU >65% of silicon die

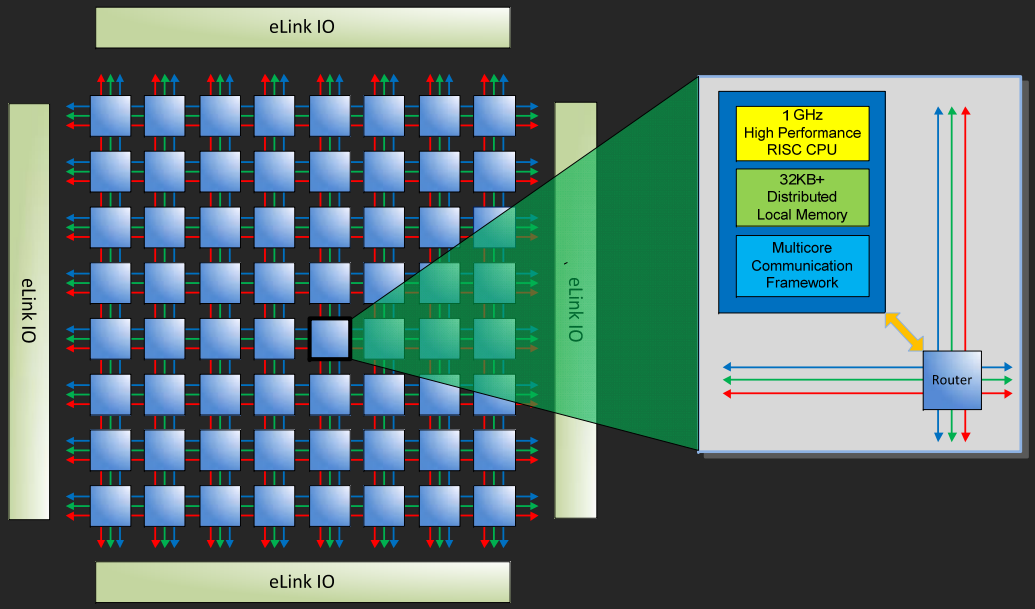
Make every transistor count

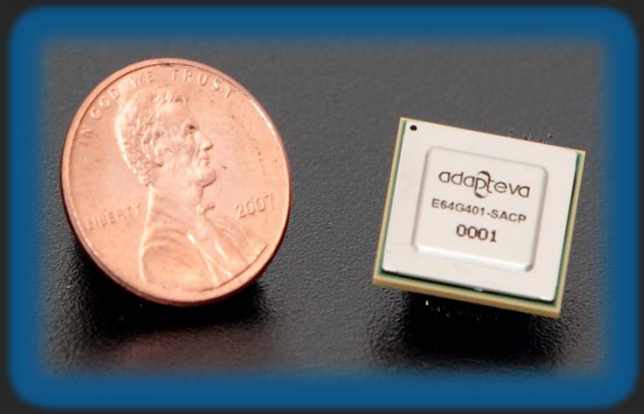
MIMD efficiency "good enough"

adapteva

Epiphany-IV Specifications

- 64 CPUs
- IEEE Floating Point (SP)
- 800 MHz Max Frequency
- 100 GFLOPS Performance
- 6.4 GB/s IO BW
- 200 GB/s peak NOC BW
- 1.6 TB/sec on chip memory BW
- 25 Billion Messages/sec
- 2MB on chip memory
- 10 mm^2 total silicon area in 28nm
- 2 Watt total chip power
- 324 ball 15x15mm BGA
- Sampling since July, 2012

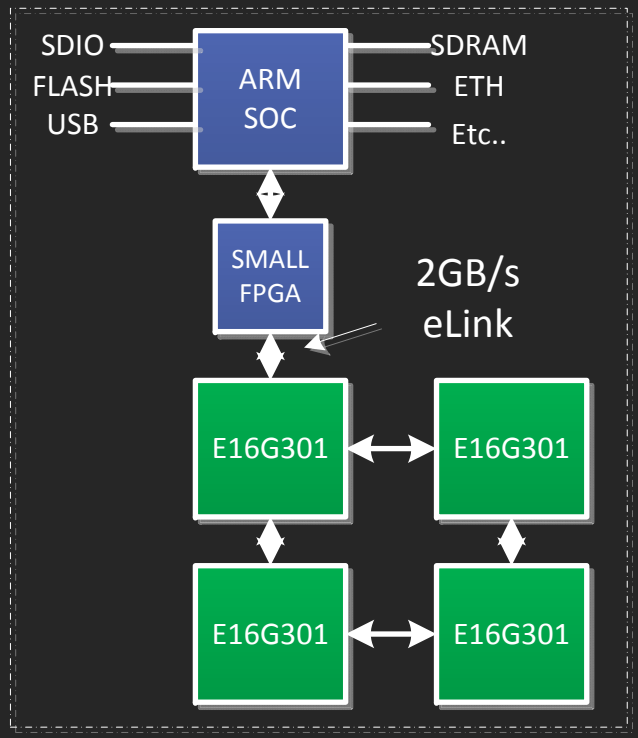


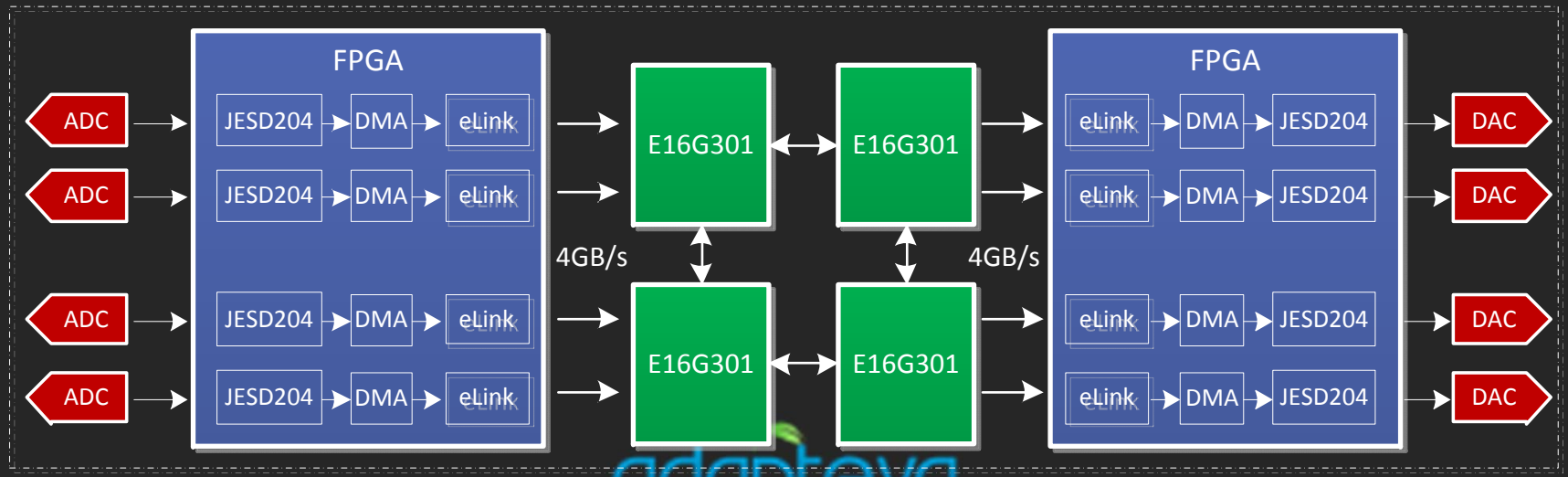


adapteva

Epiphany System Examples

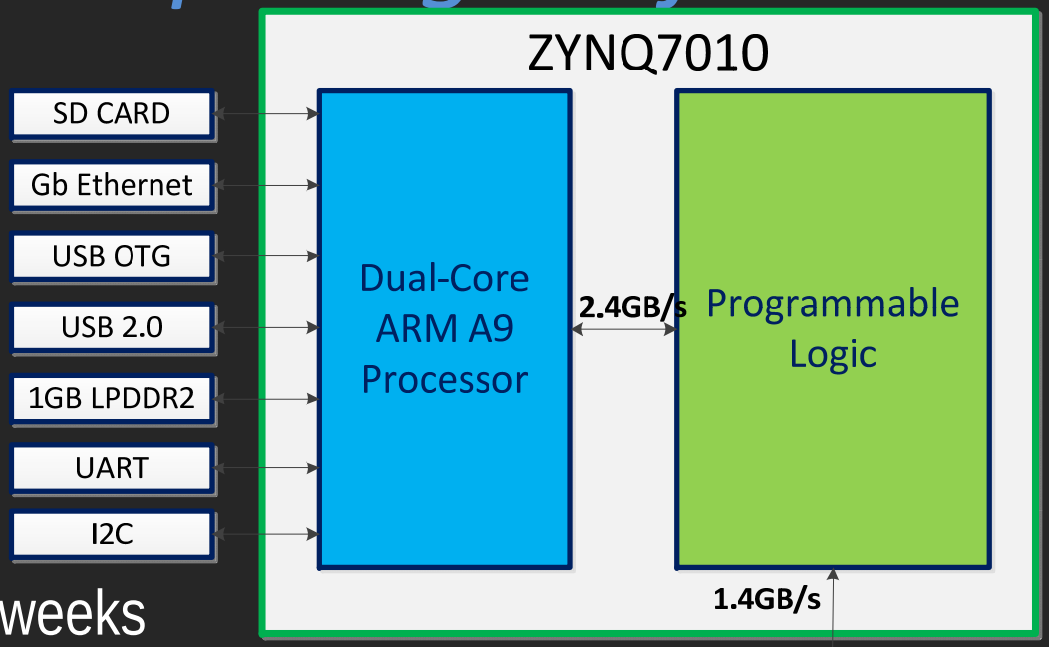
- Programmable and flexible
- Easy to develop for
- Efficient and powerful
- Scalable

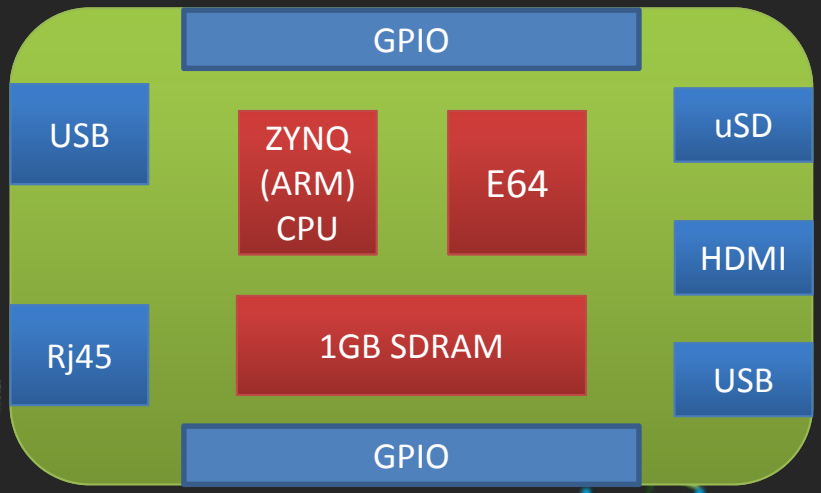


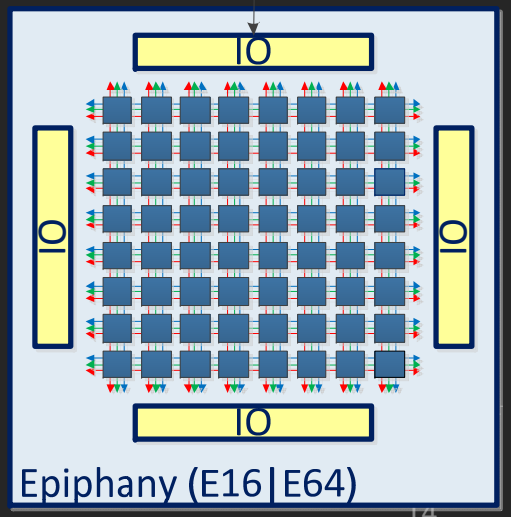


Parallella Computing Project

- Open (and ”free”):
  - Documentation
  - Board design files
  - Drivers
  - Software Tools
- Accessible (NO NDAs!)
- $100 entry point
- ~4000 devs signed up in 4 weeks

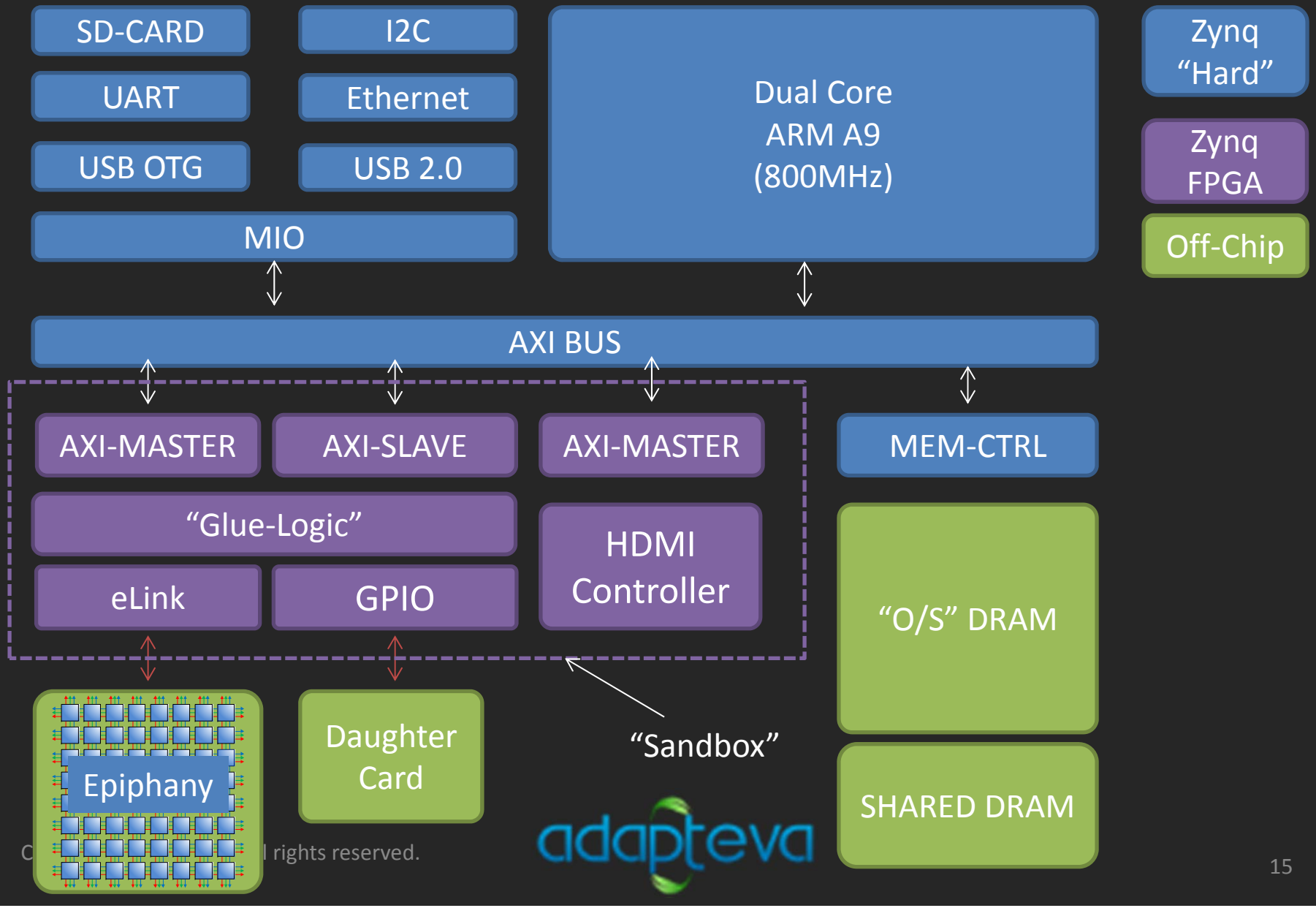






adapteva

Parallella Architecture



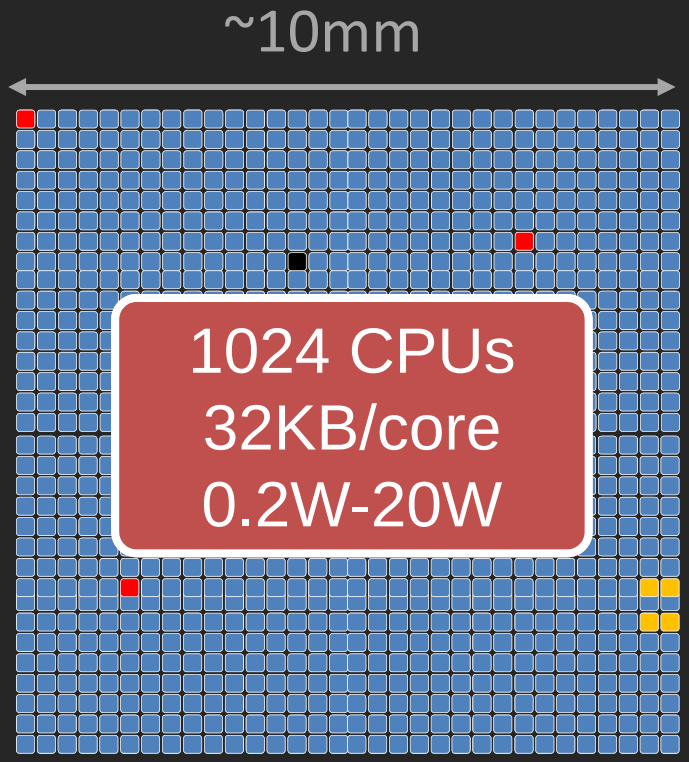
Experimental Features On The Way

- Network Traffic and Congestion Monitors
- Multicore Hardware Synchronization
- Active Message
- Multicore Breakpoint
- Hardware Loops
- Multicast Network Transactions

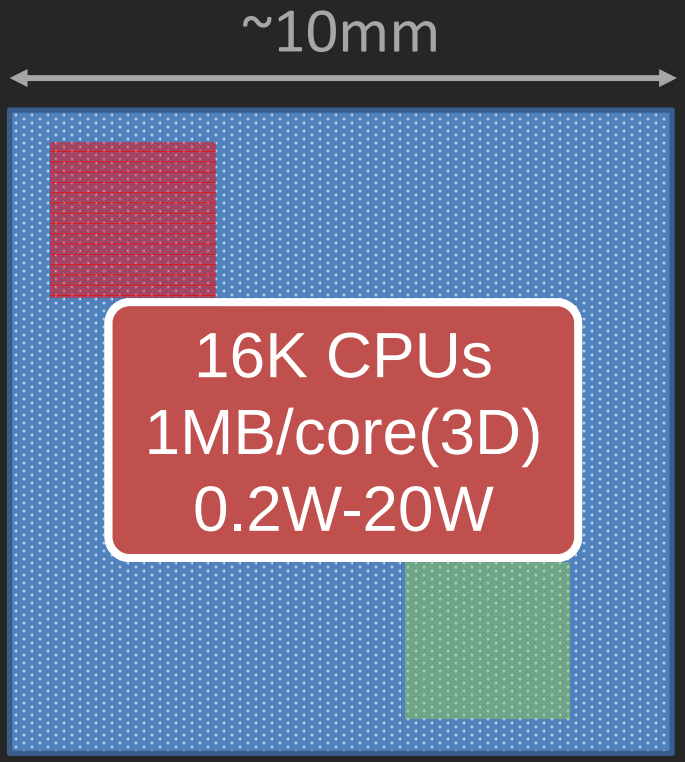
Already inside Epiphany-III/IV silicon, but need more testing!

adapteva

MIMD Manycore IS the Future!



2012



2022

adapteva