

Inventing the Future of Computing

An Alternative to GPU Acceleration in Mobile Platforms

Andreas Olofsson

andreas@adapteva.com

Linley Tech Mobile Conference 2013
April 17th & 18th, 2013 – Santa Clara, CA



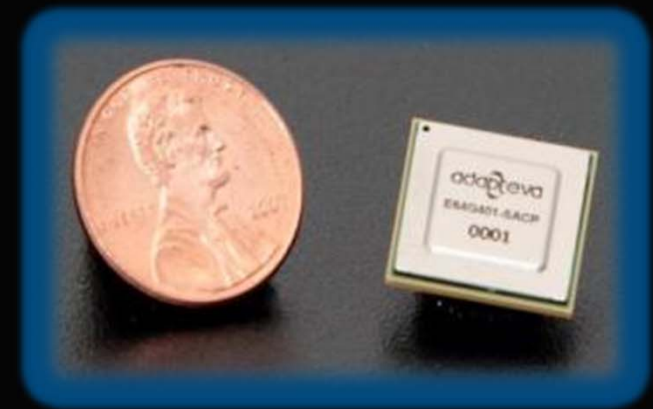
What is Adapteva

Company History:

- Fabless semiconductor company founded in 2008
- 16-core 65nm Epiphany-III chip product sampling since May 2011
- 64-core 28nm Epiphany-IV chip product sampling since July 2012
- Parallella open computing platform launched in October 2012

Notable Achievements:

- #1 in microprocessor energy efficiency
- 4 chips on \$2.5M in raised capital
- \$2M in total revenue to date
- 5K customers, 6,300 boards pre-sold
- 18 Patents pending



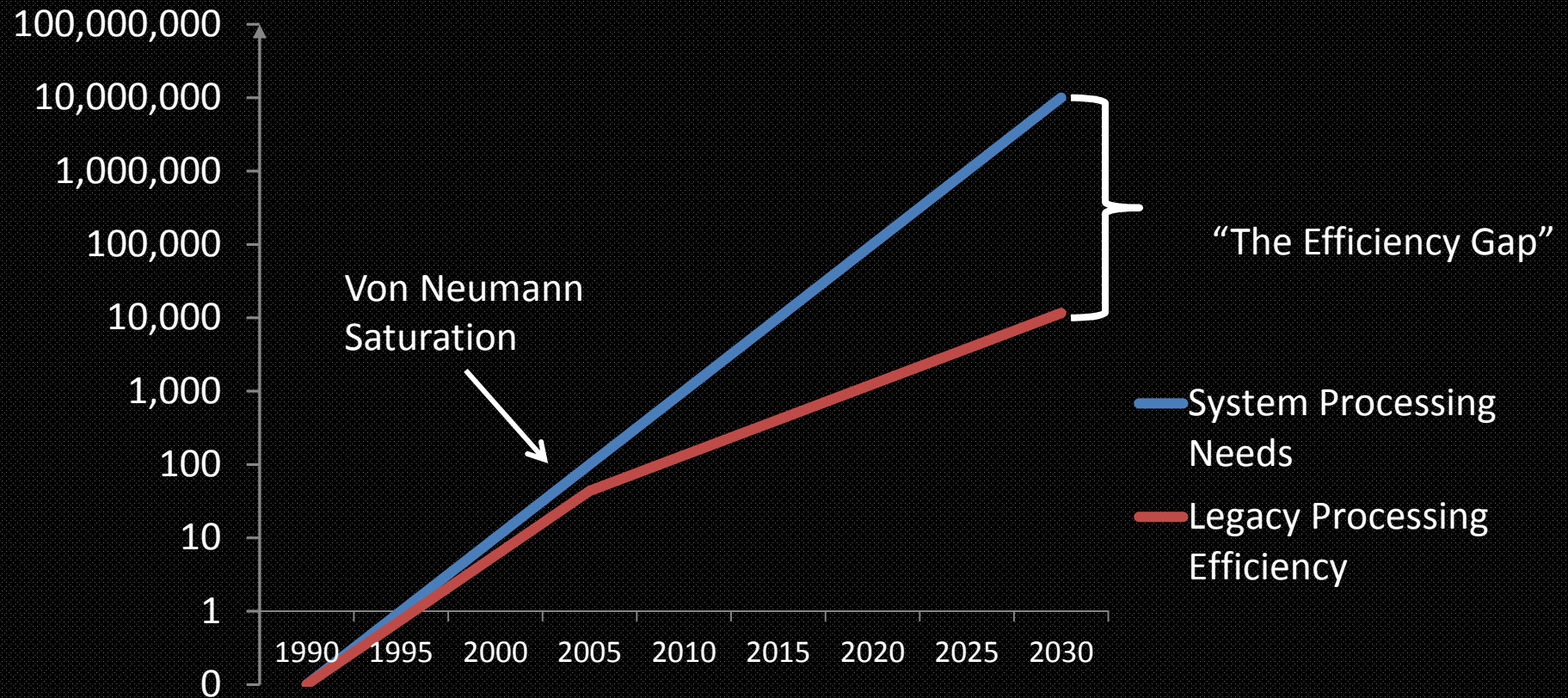
Andreas Olofsson, Adapteva

Linley Tech Mobile Conference, April 17, 2013

Copyright © Adapteva. All rights reserved.



The Mobile Energy Crisis: IT'S REAL!!



Trends that will shape the future of computing

Power Consumption

Latency Wall

Memory Bottlenecks

Yield Issues

Frequency Wall

Time to Market

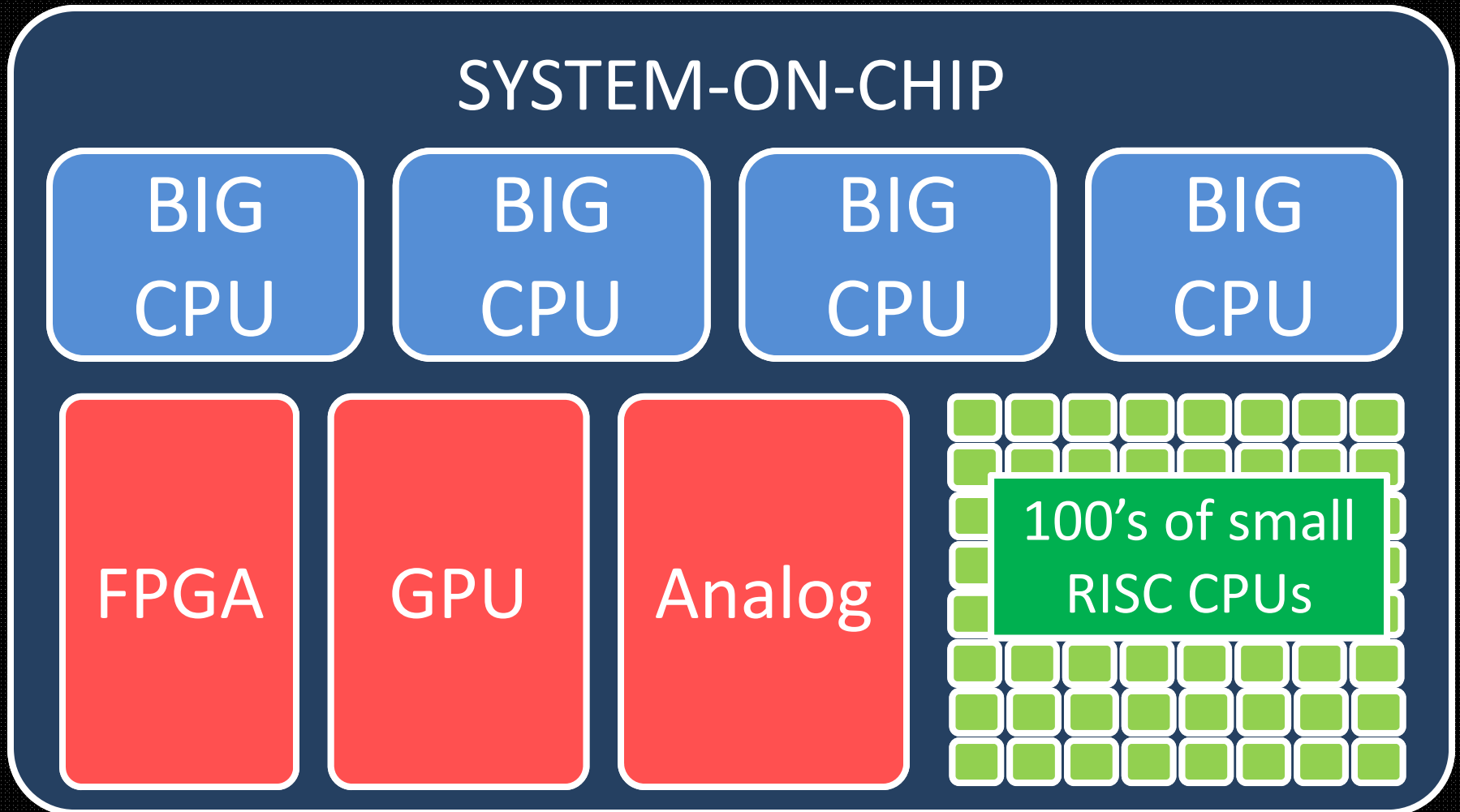
Wiring

Software Complexity

Thermal Density

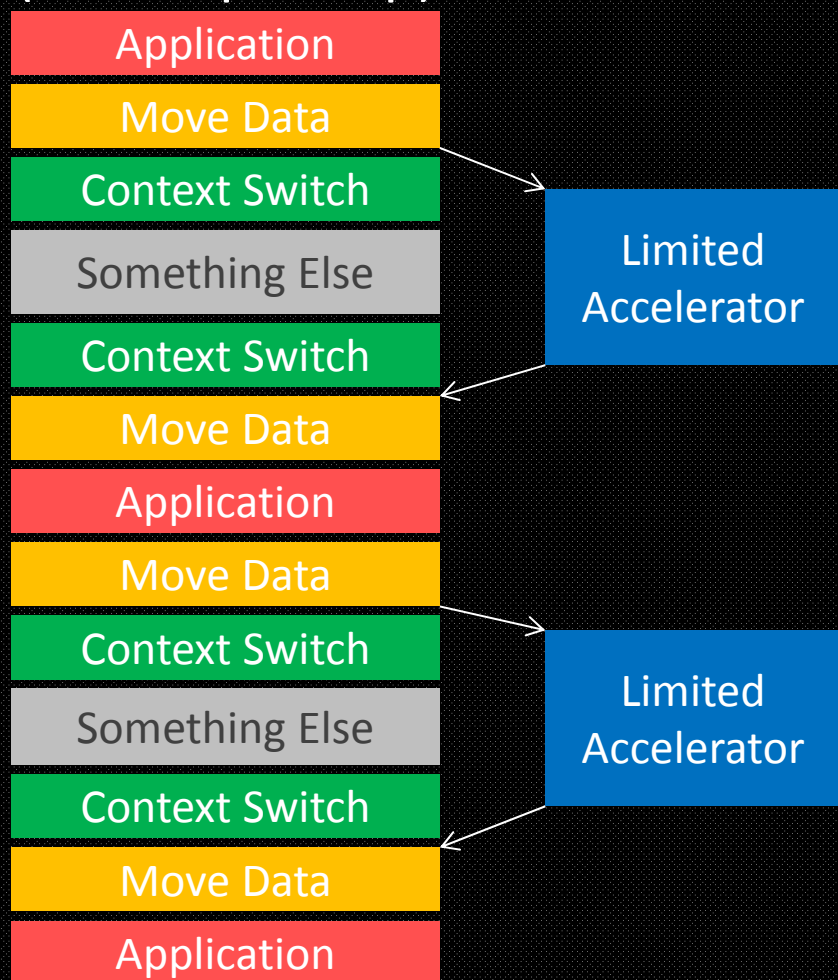
Amdahl's Law

The Solution: True Heterogeneous Computing

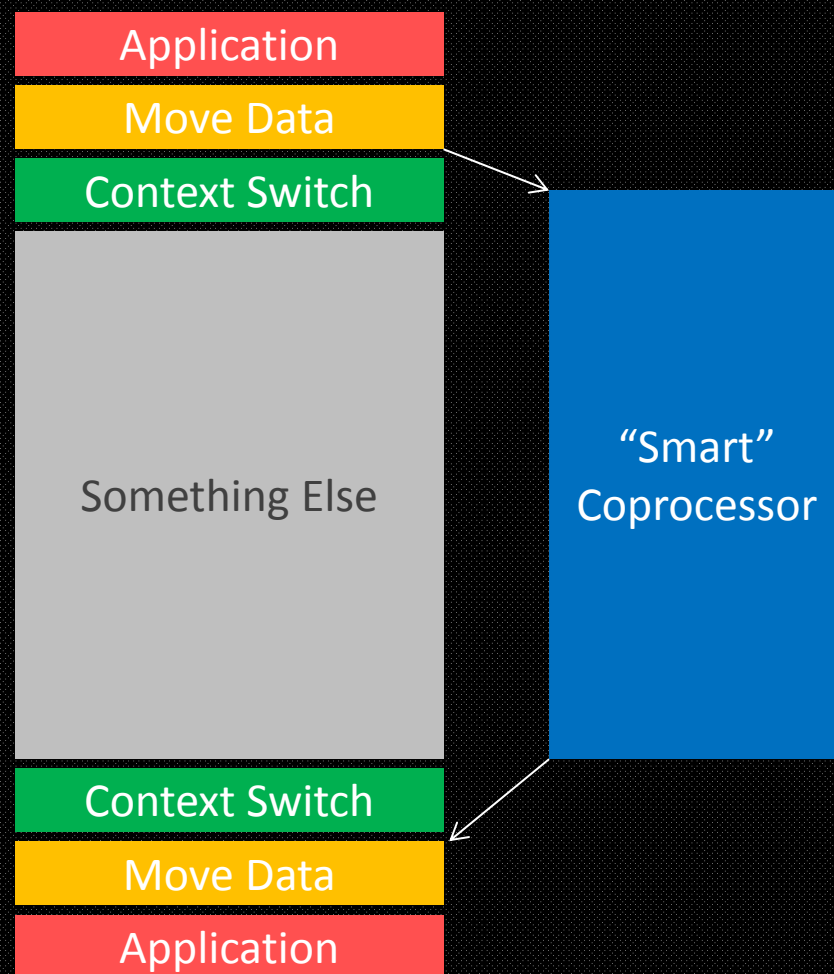


The Accelerator Challenge

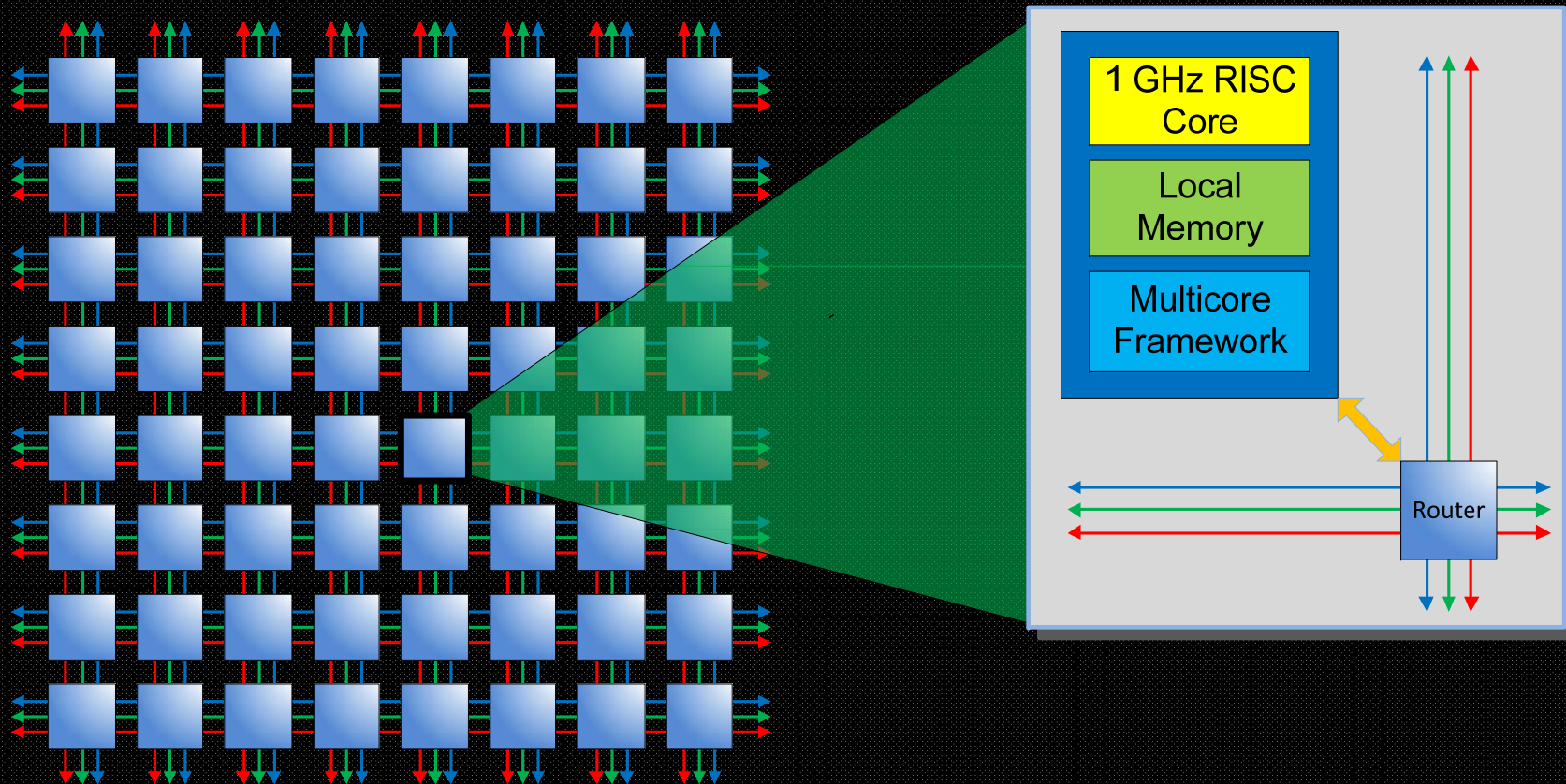
Status Quo Approach
(~1.3X speedup)



Smart Coprocessor (>10X speedup?)



Epiphany Multicore Technology



**Coprocessor for
ARM/x86/MIPS Host**

**10X Boost in Energy
Efficiency**

**True Task Parallel
Coprocessor**

Andreas Olofsson, Adapteva

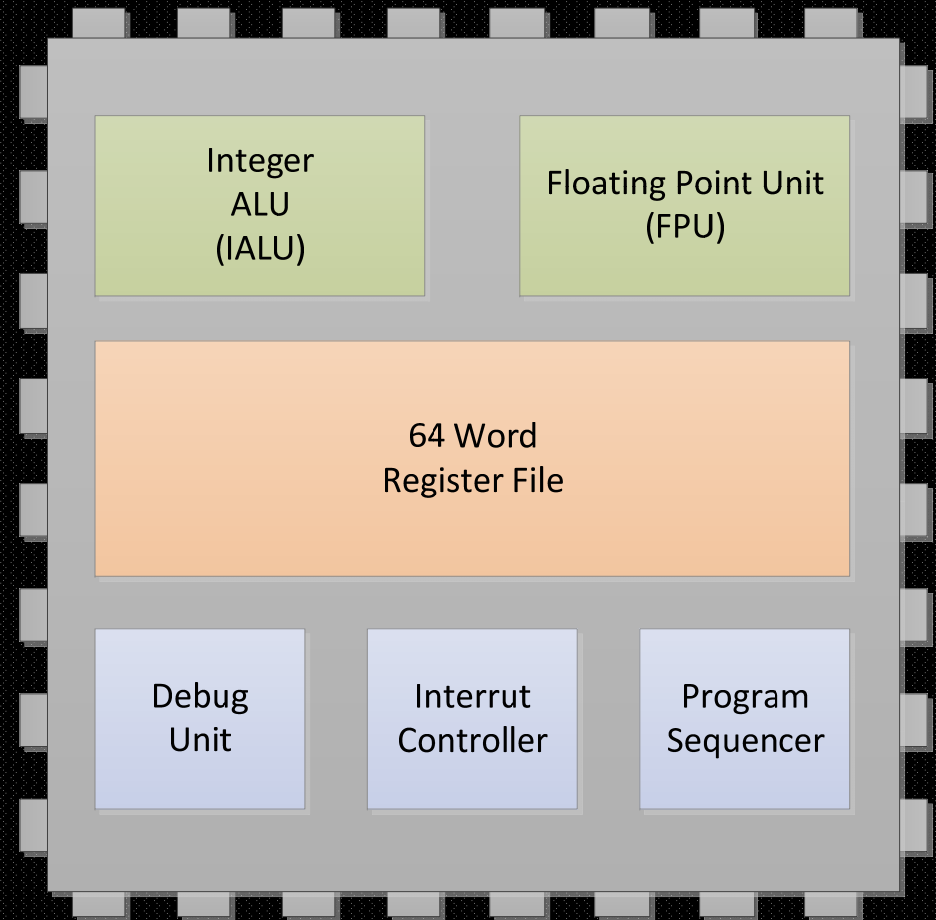
Linley Tech Mobile Conference, April 17, 2013

Copyright © Adapteva. All rights reserved.



The Epiphany Core

- 1GHz Operation at 65G
- Dual issue superscalar architecture
- 64-entry register file
- Dense instruction encoding
- Fused floating point Multiply Add
- Quad Port Local Memory
- Nested Interrupt Support
- Memory Protection
- Built in performance timers
- Built for multicore integration



Epiphany ANSI-C Benchmarks

(Cycles)	Naïve C	Optimal C	Theoretical	C-Efficiency
8x8 Matrix Multiplication	2852	773	512	66%
16 Tap FIR Filter (32 points)	1562	620	512	82%
Bi-quad IRR Filter (32 points)	n/a	991	768	77%
Dot-product (256 point)	800	557	256	49%

1 day per benchmark
(compare to GPUs?)

	Adapteva E64 800 MHz	Tilera GX36 1.4GHz	Intel Xeon L5640 2.2GHz	Nvidia Tegra-2 1GHz
CoreMark TM Score	77,912	165,276	118,571	5,866
# Cores	64	36	8	2
Power	2W	~30-50W	~50-100W	~1-2W
1024-Core Chip	2,493,184	n/a	n/a	n/a

Server Level Performance at 2Watts!!

Andreas Olofsson, Adapteva

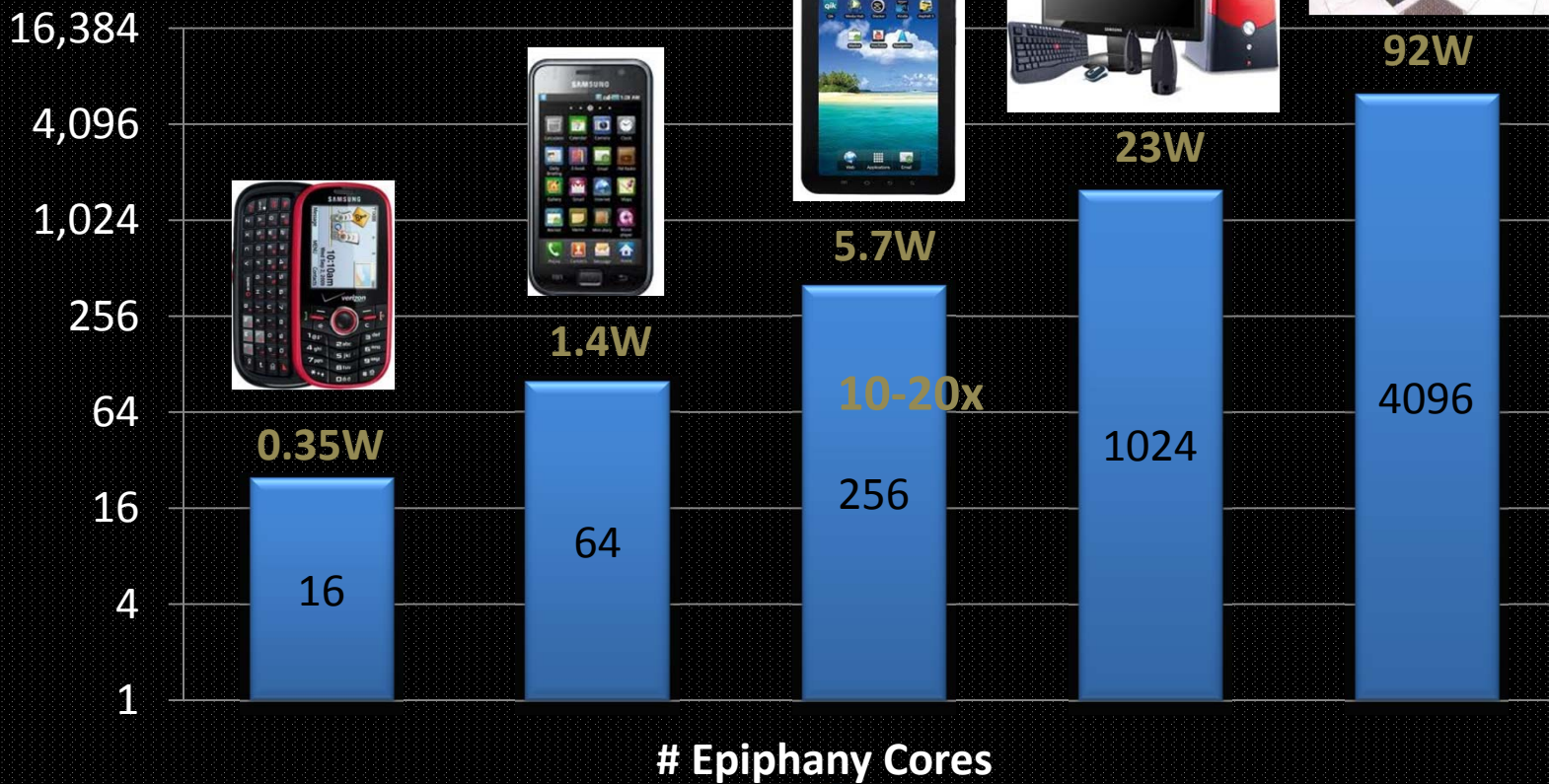
Linley Tech Mobile Conference, April 17, 2013

Copyright © Adapteva. All rights reserved.



Epiphany: A Truly Scalable Architecture

A Single Unified Instruction Set Architecture!



Andreas Olofsson, Adapteva
Linley Tech Mobile Conference, April 17, 2013
Copyright © Adapteva. All rights reserved.

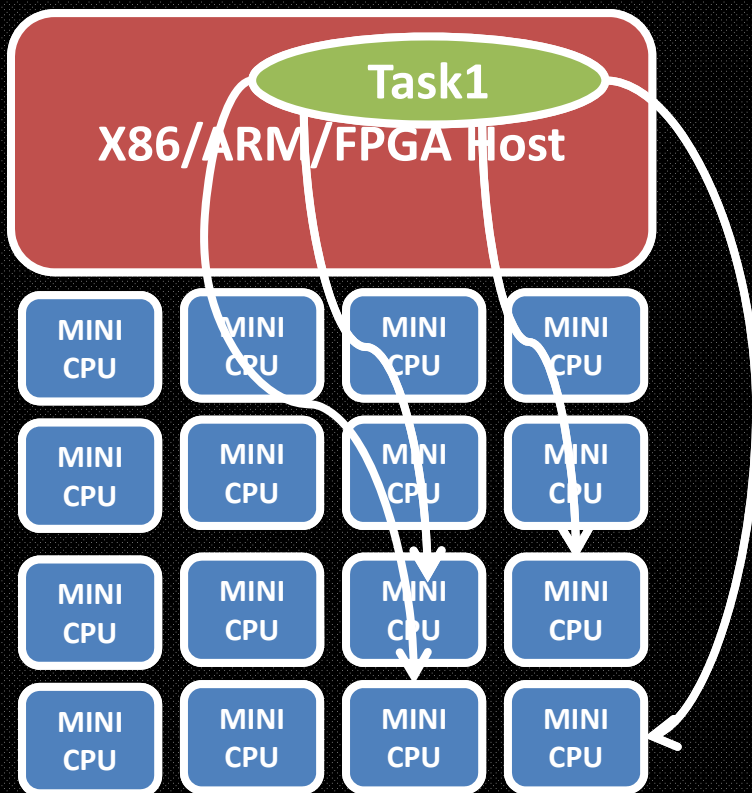


Epiphany Programming Models

MODEL #1

DATA PARALLEL MODEL

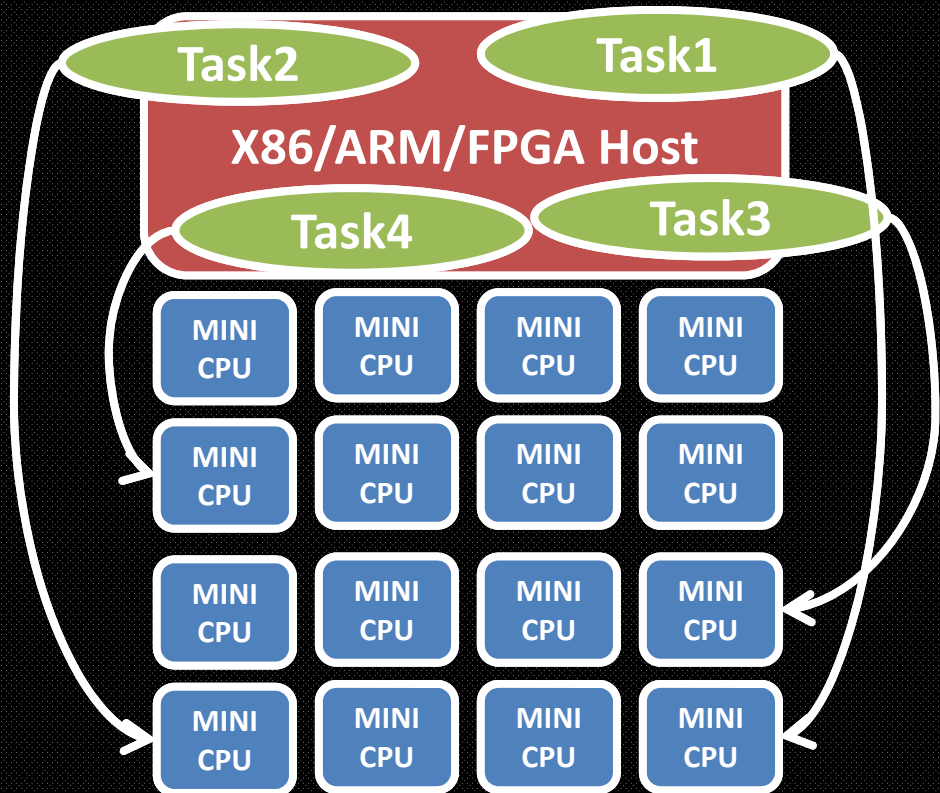
- openCL programmable
- Easy integration with C/C++
- openMP/MPI roadmap



MODEL #2

WORKER BEE MODEL

- Great for up to 2GFLOPS
- Supports standard C/C++
- "Cloud on a chip"



Andreas Olofsson, Adapteva

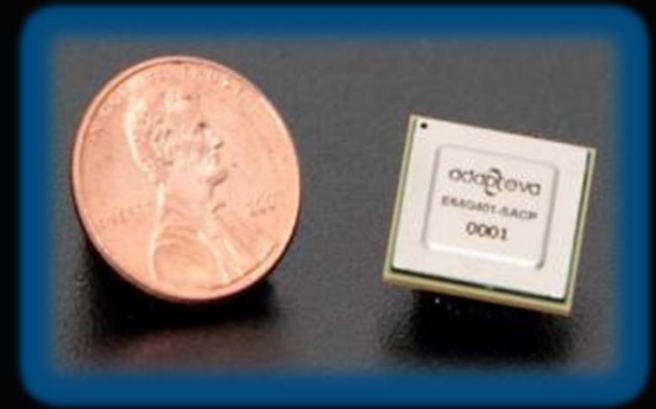
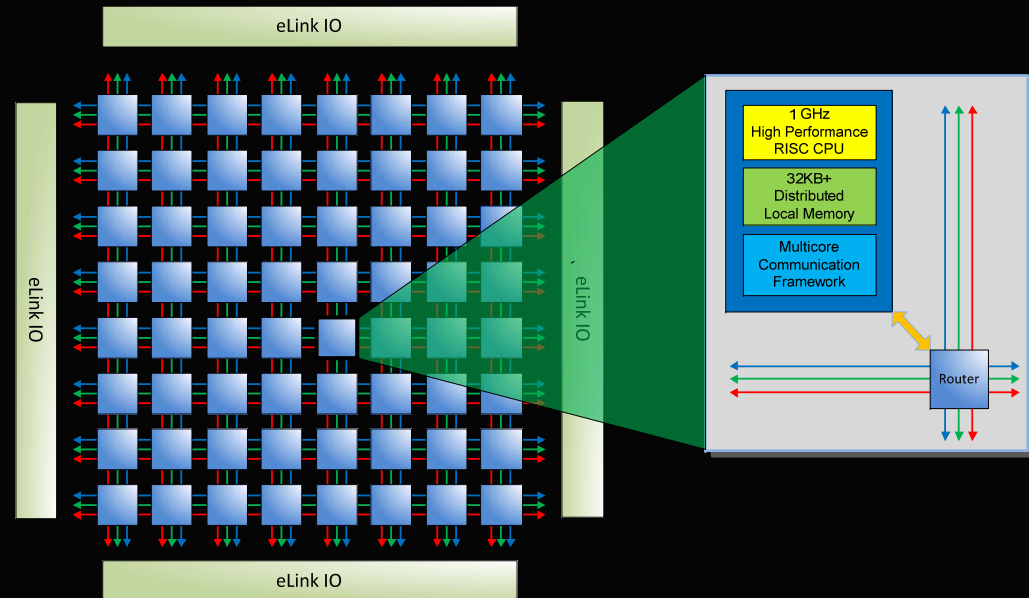
Linley Tech Mobile Conference, April 17, 2013

Copyright © Adapteva. All rights reserved.



Epiphany-IV Specifications (28nm)

- 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 mm² total silicon area in 28nm
- 2 Watt total chip power
- 324 ball 15x15mm BGA
- Sampling since July, 2012



Andreas Olofsson, Adapteva

Linley Tech Mobile Conference, April 17, 2013

Copyright © Adapteva. All rights reserved.



Architecture Comparison

Technology	FPGA	DSP	GPU	CPU	Epiphany
Process	28nm	40nm	28nm	32nm	28nm
Programming	VHDL	OCL/C++/C	CUDA/OCL	OCL/C/C++	OCL/C/C++
Area (mm²)	590	108	294	216	10
Chip Power (W)	40	22	135	130	2
"CPUs"	n/a	8	32	4	64
Max GFLOPS	1500	160	3000	115	102
GHz * Cores	n/a	12	32	14.4	51.2
Compile Time	Hours	Minutes	Minutes	Minutes	Minutes
L1 Memory	6MB	512KB	2.5MB	256KB	2MB

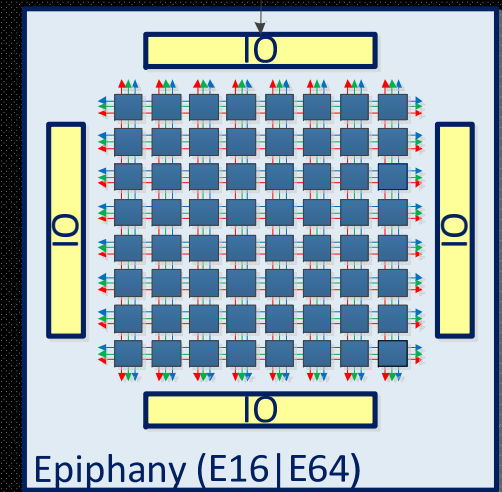
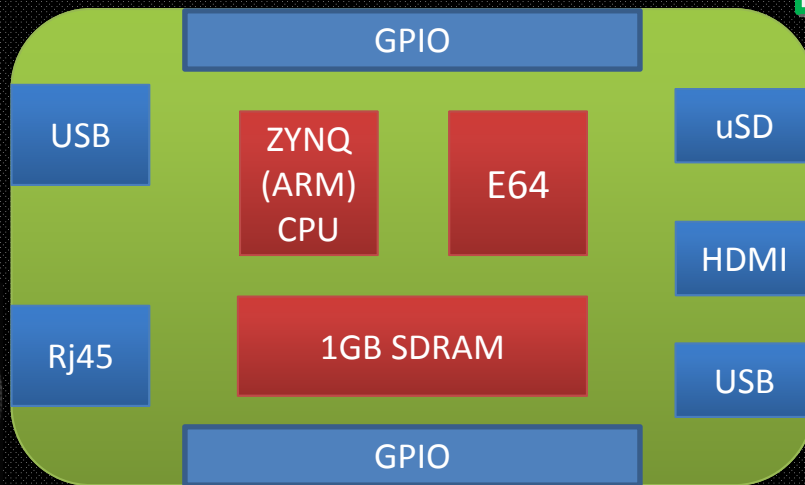
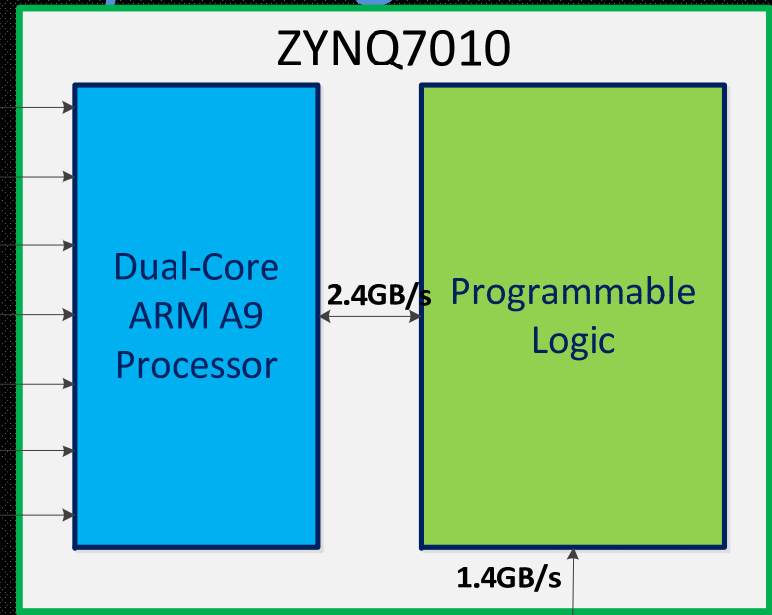
Efficiency is everything

Peak performance means very little

No magic bullet!

Parallella: (Super) Computing Boards

- OPEN! (and free):
 - Documentation
 - Board design files
 - Drivers
 - Software Tools
- Accessible (No NDAs!)
- \$100 entry point



Andreas Olofsson, Adapteva
Linley Tech Mobile Conference, April 17, 2013
Copyright © Adapteva. All rights reserved.



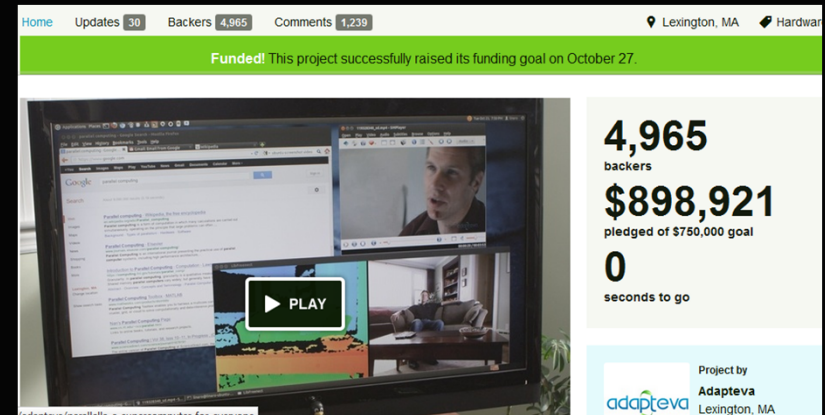
Kickstarter Campaign Stats



- 5,000 customers
- 6,300 boards "pre-sold" in 4 weeks
- 67 countries, all 50 US states
- 50-75% of backers are developers
- 4,000 more signups since Jan 1st

Customer Application Classes:

- SDR
- Ray tracing/rendering
- Image processing
- Robotics
- Gaming



- Cryptography
- Media Server
- Distributed Computing
- Signal processing
- HPC

Andreas Olofsson, Adapteva

Linley Tech Mobile Conference, April 17, 2013

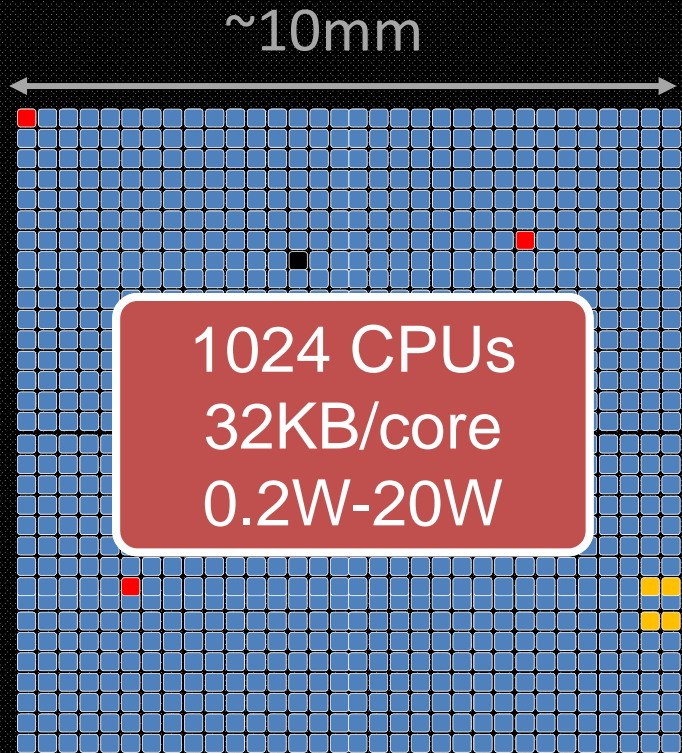
Copyright © Adapteva. All rights reserved.



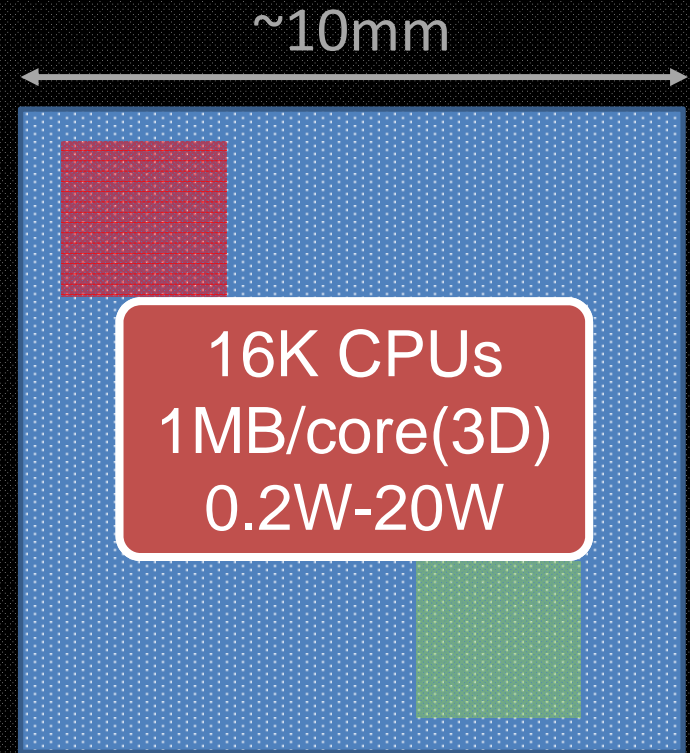
When we say open, we mean it

- [Epiphany Architecture Reference Manual \(PDF\)](#)
- [Epiphany SDK Reference Manual \(PDF\)](#)
- [Epiphany-III Datasheet \(PDF\)](#)
- [Parallella Reference Manual \(PDF\)](#)
- [Face Detection using the Epiphany Multicore Processor](#) (source code)
- [Using a Scalable Parallel 2D FFT for Image Enhancement](#) (source code)
- [Scalable Parallel Multiplication of Big Matrices](#) (source code)
- [Epiphany Drivers](#) (source code)

We're just getting started!



2012



2022

Conclusion

The Future of Computing is:

- Efficient
- Heterogeneous
- Open platforms
- Task parallel

Industry Challenges Ahead:

- Rebuild the computer ecosystem
- Rewrite billions of lines of code
- Educate millions of programmers
- Rewrite the education curriculum