

---

# MPSoC 2004

**Business Challenges and MP**

**Mike Muller**

**CTO**

**ARM**



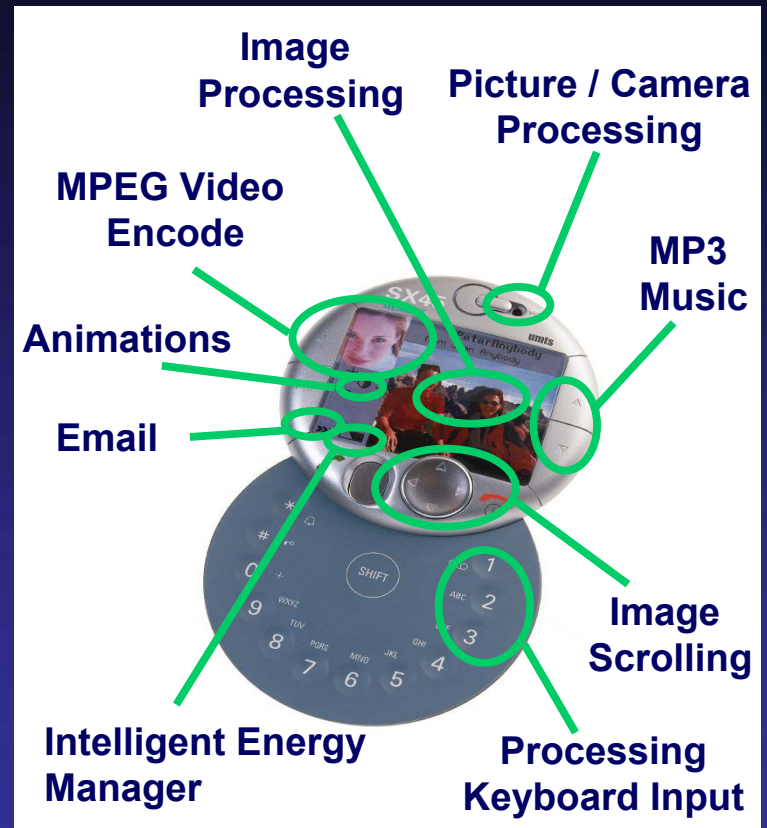
# Is IP Just a Passing Fashion?

---

- In 1983
  - Transistors vs. Gates
  - Transparent Latches vs. D-Types
  - 45° vs. 90° Manhattan Rules
  - Uncommitted Logic Arrays vs. ASICs
- In 1993
  - Gates
  - D-Types
  - 90° Manhattan Rules
  - ASICs
- In 2003
  - Back to the beginning

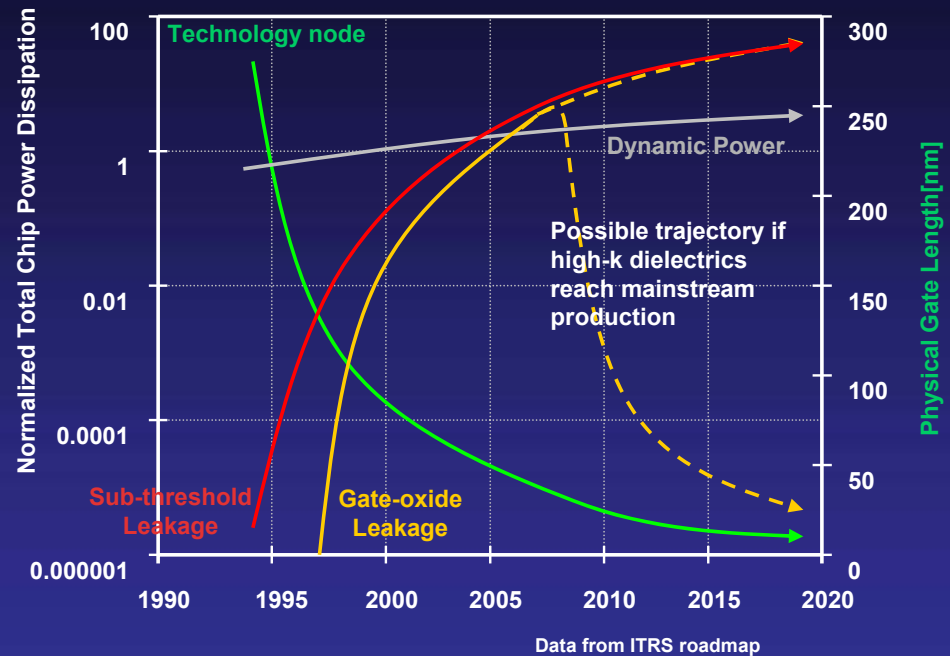
# Need for Parallel Processing

- Run multiple applications simultaneously
- Execute multiple algorithms
  - Decode & encode both video and audio stream
- Uniprocessor Solutions
  - SIMD
  - Media signal processing
  - Deeper pipelines
  - Multi-issue
- Multiprocessor solutions



# Is Silicon Our Friend?

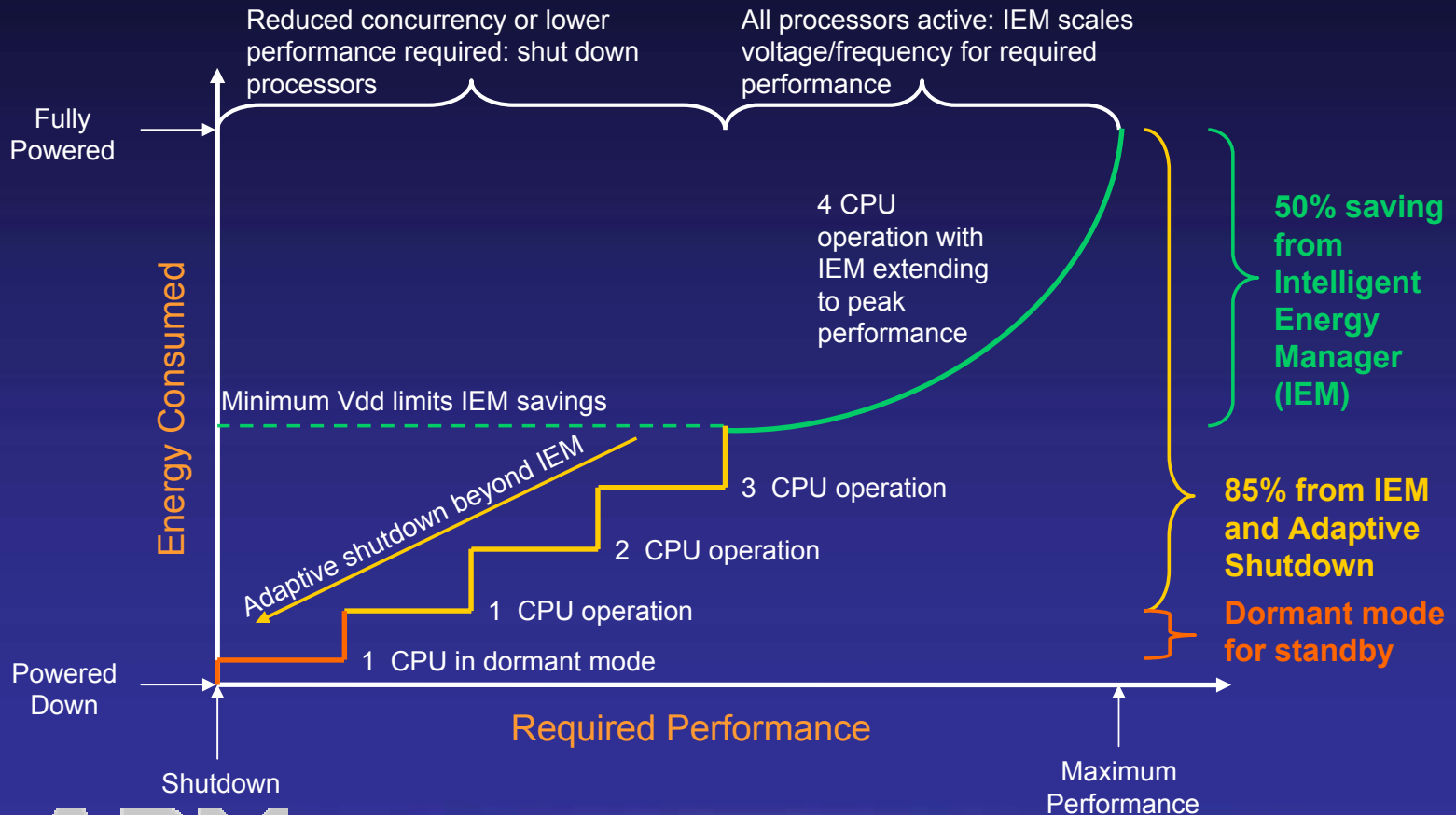
- Technology scaling trends are not in our favour
  - Leakage power
  - New processes are expensive
  - Diminishing performance gains from process scaling
  - Dynamic power remains high



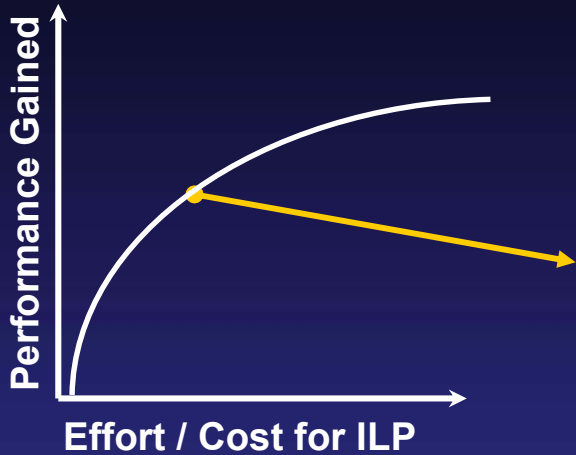
- Solutions need to cut across traditional boundaries
  - (SW / architecture / microarch / circuits)

# MPCore Multiprocessor: Saving Energy

MPCore multiprocessor extends control over power usage by providing both voltage and frequency scaling and turning off unused processors

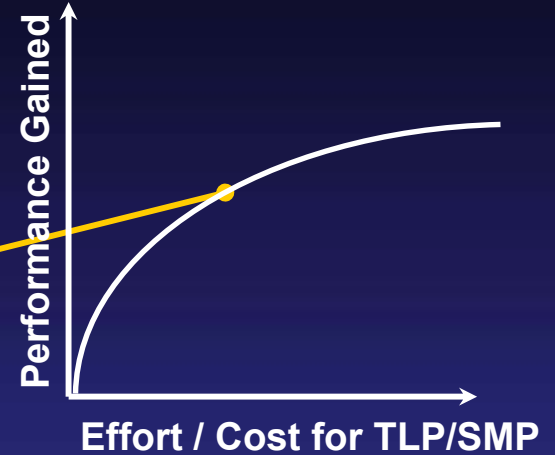
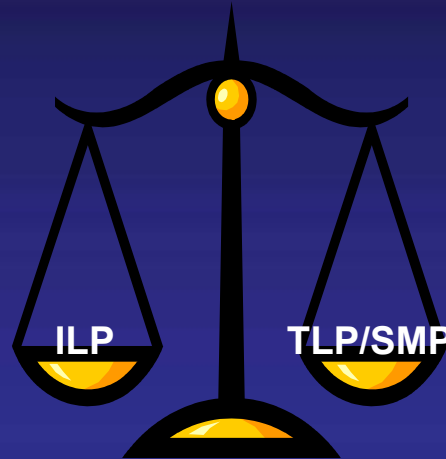


# Where Best to Apply Parallelism



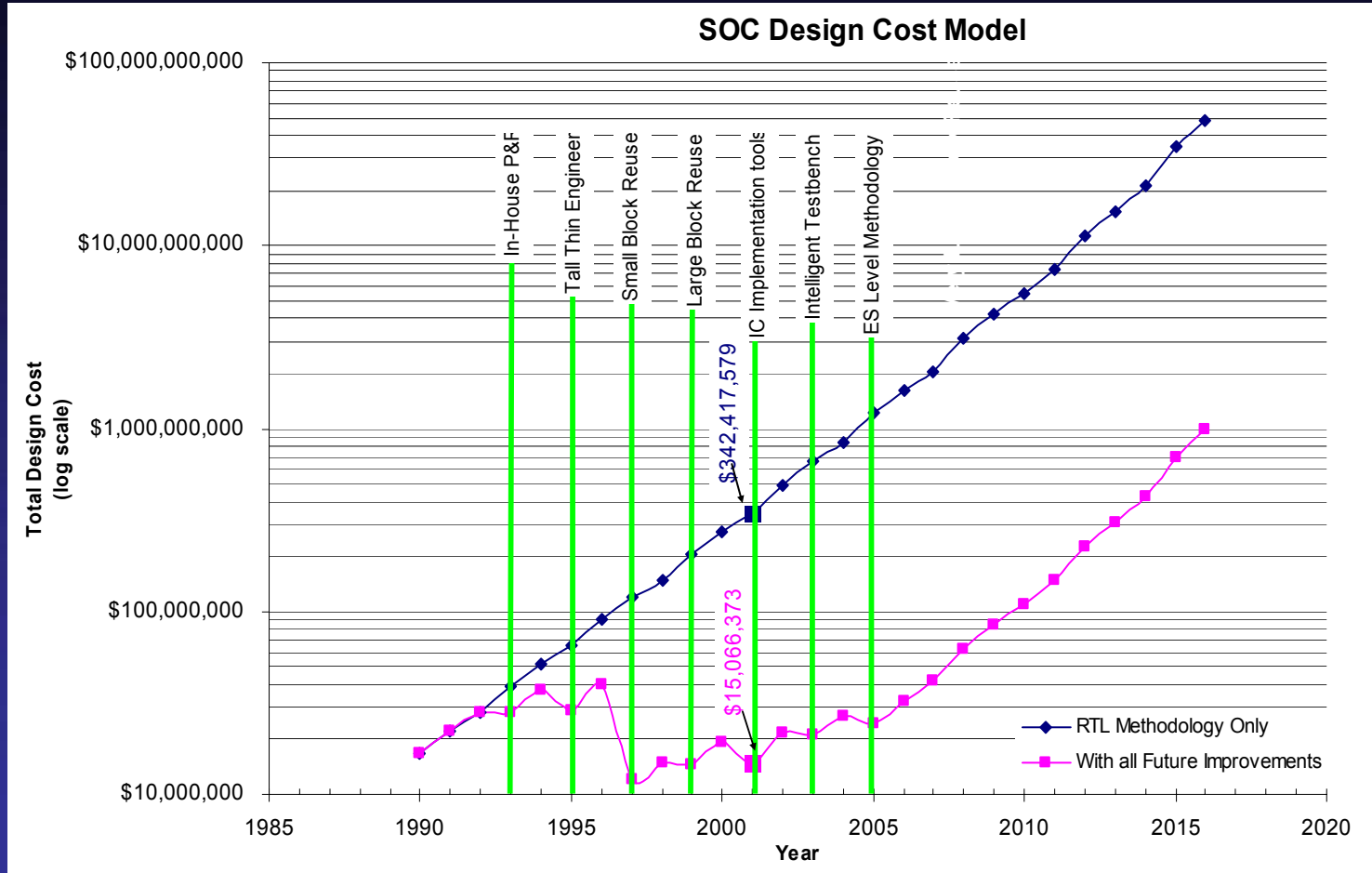
Expensive on hardware to continue to attempt to extract instruction level parallelism (ILP)

'Optimal' support for both ILP and TLP brings the most performance for the least cost / effort



Expensive on hardware and software to continue to attempt to share work between more instances of thread level parallelism (TLP)

# The Design Crisis is here to stay



ITRS 2003 SoC Design Cost Model/Synopsys



# IP Will Bring 75% of Productivity Gains

Year	DT Productivity Impact	Gates Per Designer per Year	Description
1990	Design Technology	4,000	
1993	In-House Place and Route	5,550	Automated Block Placement and Routing
1995	Tall Thin Engineer	9,090	Engineer can pursue all tasks to complete design block from RTL to GDSII
1997	Small (2K-75K) Block Reuse	40,000	Blocks from 2,500 –74,999 gates
1999	Large (75K-1M) Block Reuse	56,000	Blocks from 75,000-1M gates
2001	IC Implementation Suite	91,000	Tightly integrated toolset that goes from RTL synthesis to GDSIII through IC place and route
2003	Intelligent Test Bench	125,000	RTL verification tool (cockpit) using an ES-level description with tracking and reporting of code coverage
2005	ES Level Methodology	200,000	A behavioral and an architectural level (where HW and SW are identified and handed off to design teams)
2007	Very Large (>1M) Block Reuse	600,000	Blocks >1M gates; intellectual-property cores

ITRS 2003 SoC Design Cost Model/Synopsys

- Multi Processing is excellent IP Reuse

THE ARCHITECTURE FOR THE DIGITAL WORLD™



# Major Internal Costs in CPU Development

1

## New Instructions Definition

- Upfront definition of instructions to be added to ISA

2

## CPU Implementation

- CPU design, implementation, verification and validation
- Software tools, compilers, debuggers and boards

3

## Platform Implementation

- Platform infrastructure
- Busing
- Peripheral development
- Implementation, verification and validation
- OS porting

4

## Ongoing Support Cost (Sustaining)

- In-house support staff

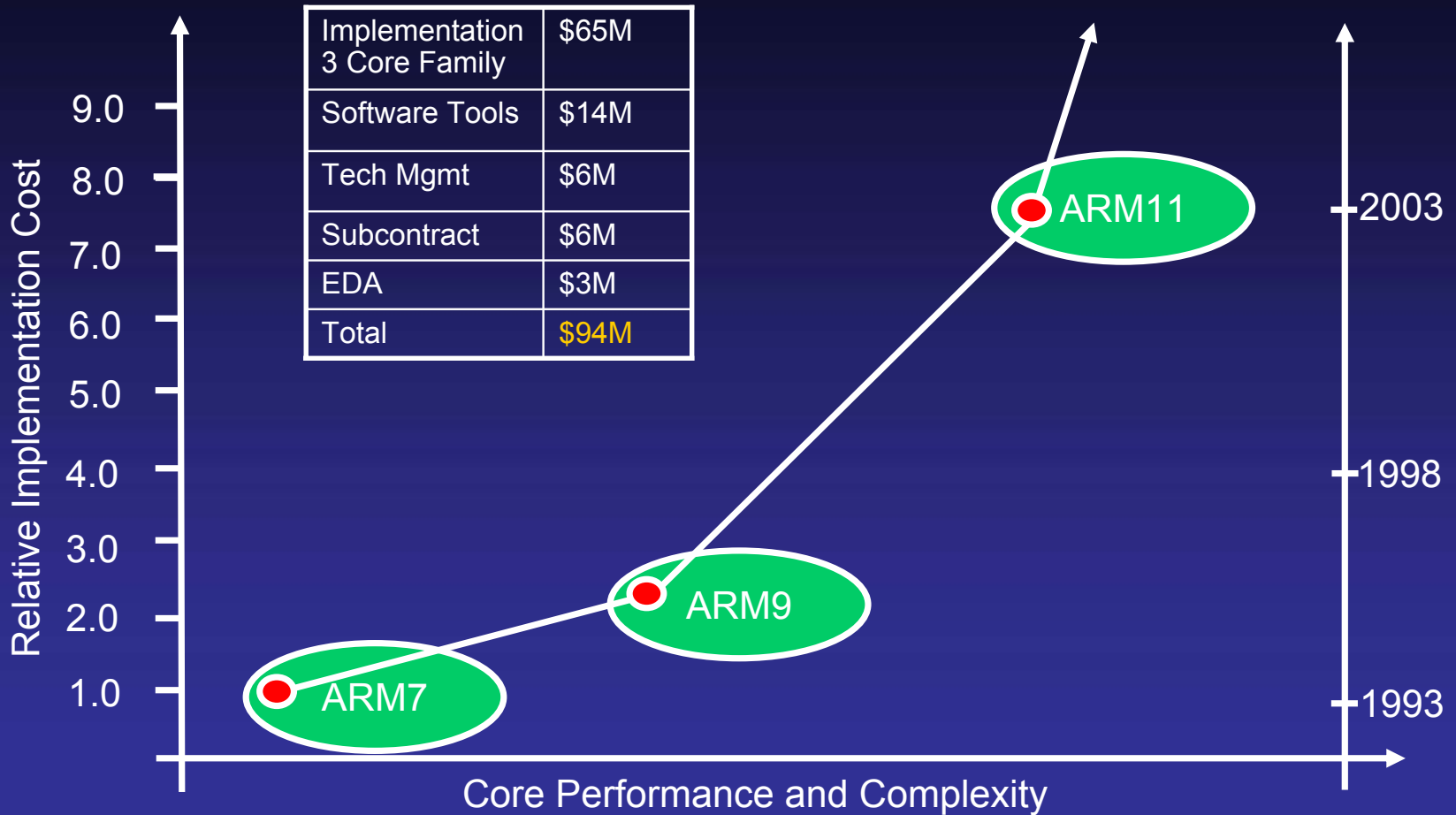
# ① Cost of New Instruction Definition

---

- Architectural definition of new instructions
- Mostly research with some high-level engineering
- Builds on cumulative dollars already invested in the architecture

Addition of significant ISA extension	\$3M
---------------------------------------	------

## ② Incremental Implementation Cost



## ③ Cost of SoC Platforms

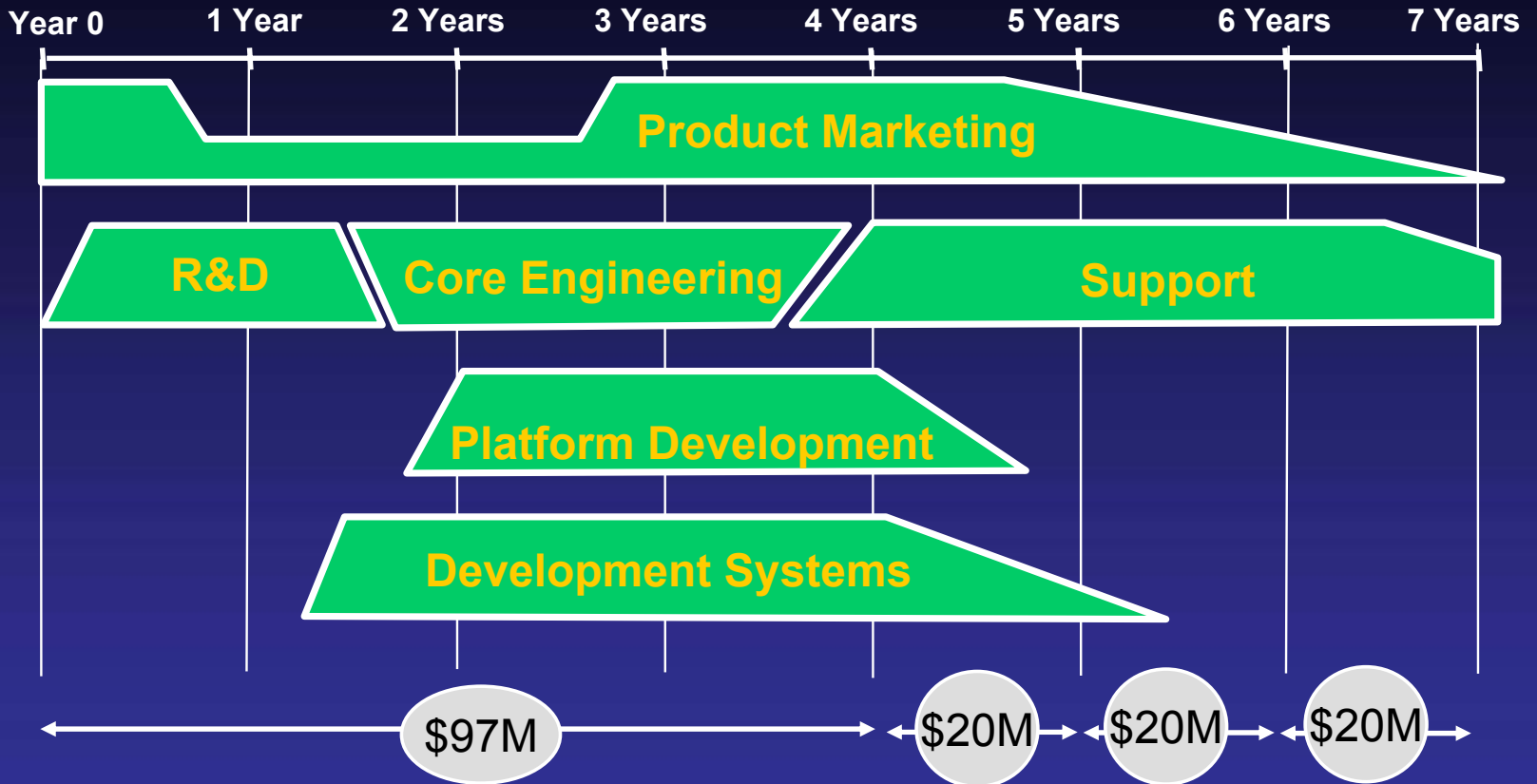
---

- Platforms
  - \$10-15M per platform
- AMBA bus infrastructure
  - Ongoing cross-platform cost
  - Total ~\$6M per generation
  - >\$20M invested to date
- Peripheral development
  - Ongoing cross-platform cost
  - \$3M per year

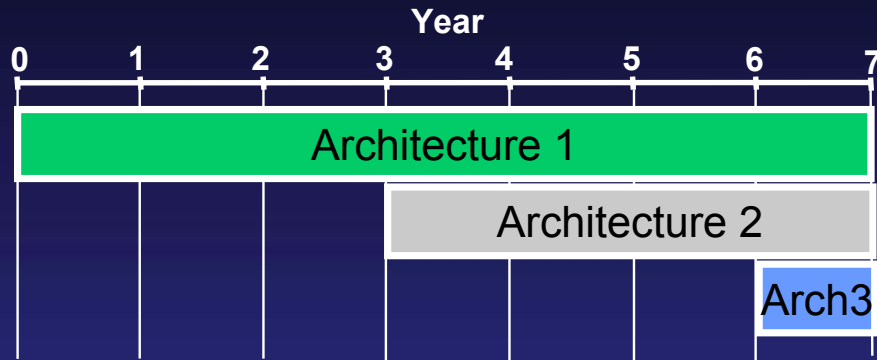
## ④ Ongoing Sustaining Costs

	Cost	
10 Support Heads	\$3M	
3 Product Managers	\$1M	
AMBA	\$2M	On-chip Bus Development
Peripheral Engineering	\$3M	Platform Peripherals
3 <sup>rd</sup> Party Tools Community	\$8M	Apportioned Share of Spend
Brand and High-Level Mkting	\$3M	Apportioned Share
<b>Total</b>	<b>\$20M/yr</b>	

# 5-Year Time to Revenue

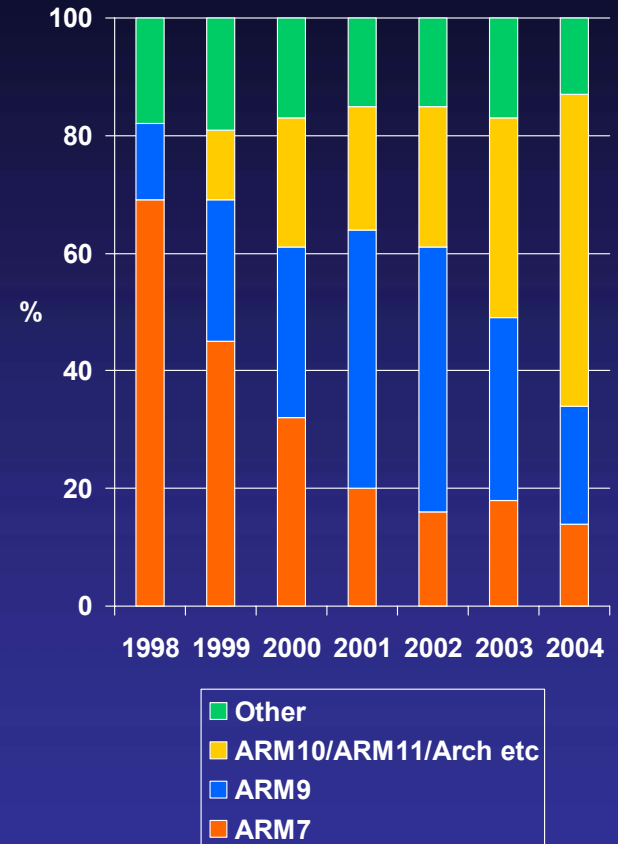


# Commitment to Roadmap Treadmill

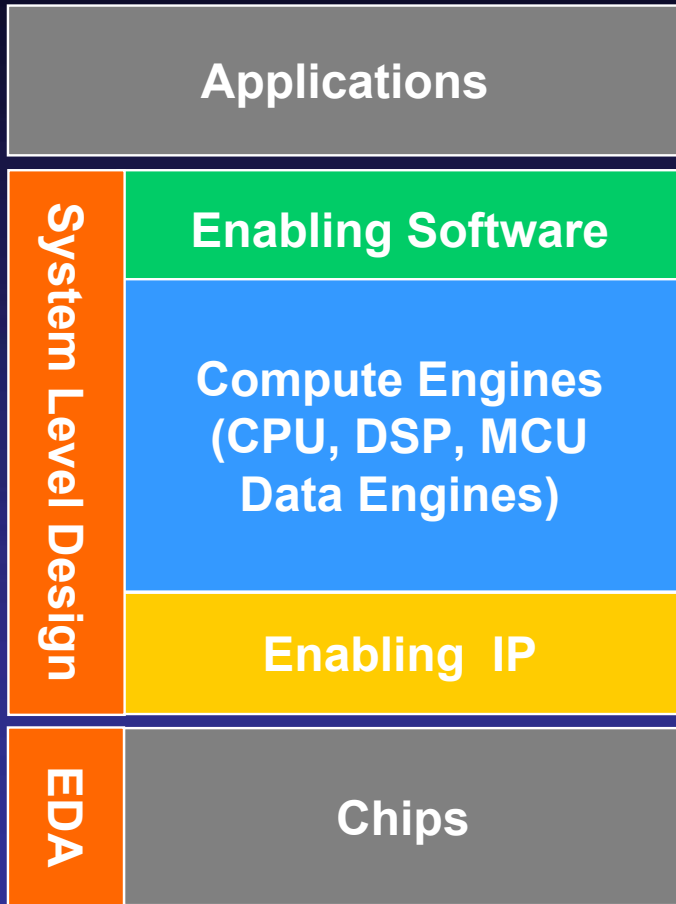


- Each generation is more expensive
- BUT it's not just a top right world
- Need to invest across the performance spectrum
- Need to reuse for MP solutions

Revenue by Family



# So is the IP Space Viable?



Yes,  
at 1 Billion Cores  
a Year



# ARM

THE ARCHITECTURE FOR THE DIGITAL WORLD™



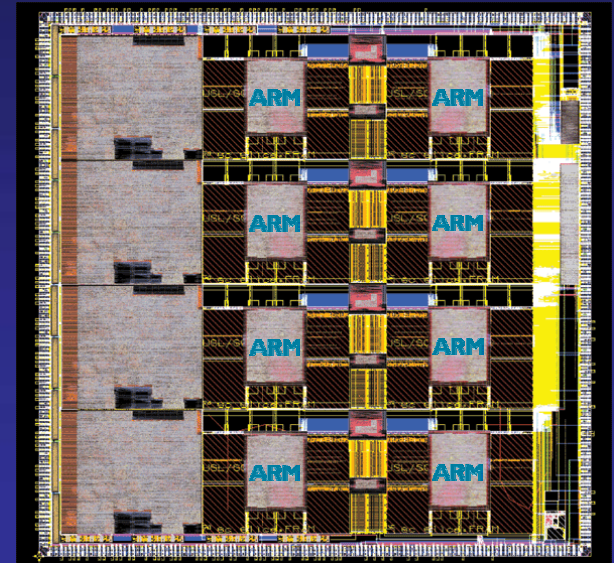
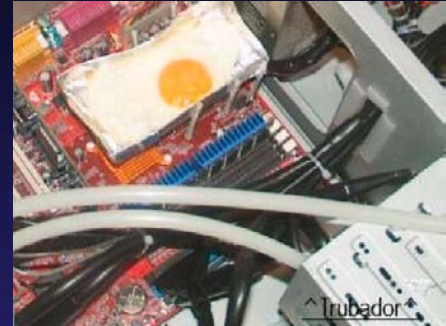
# But It's Expensive

---

- Engineering
  - Annual cost of the CPU treadmill +\$45M
  - Cost of a current high-end architecture +\$97M
  - Next generation x 4 again
- Third-party community
  - 10 x the investment
- Time to revenue ~5 years
  - 2x the life expectancy of CEO
  - A little better for a CTO 😊
- As implementation costs rise it's even more important to have standard platforms

# Technology Challenges

- 1-billion transistor ASIC by 2007
  - For example 20 x 1GHz CPUs
- Critical Issues
  - Programming model
  - Power & noise
  - Reliability and security
  - Correctness & verification
  - Interconnect



8 x ARM966E-S™ from Agere

# A Quick Way to 1 Billion in 3D

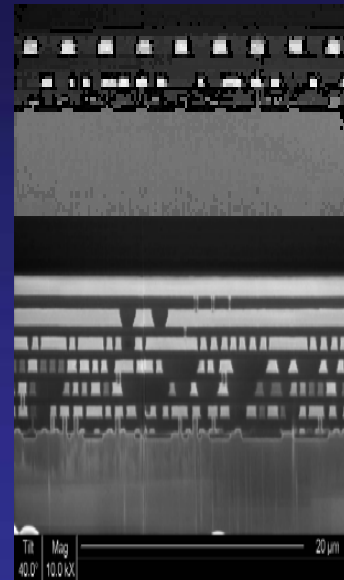
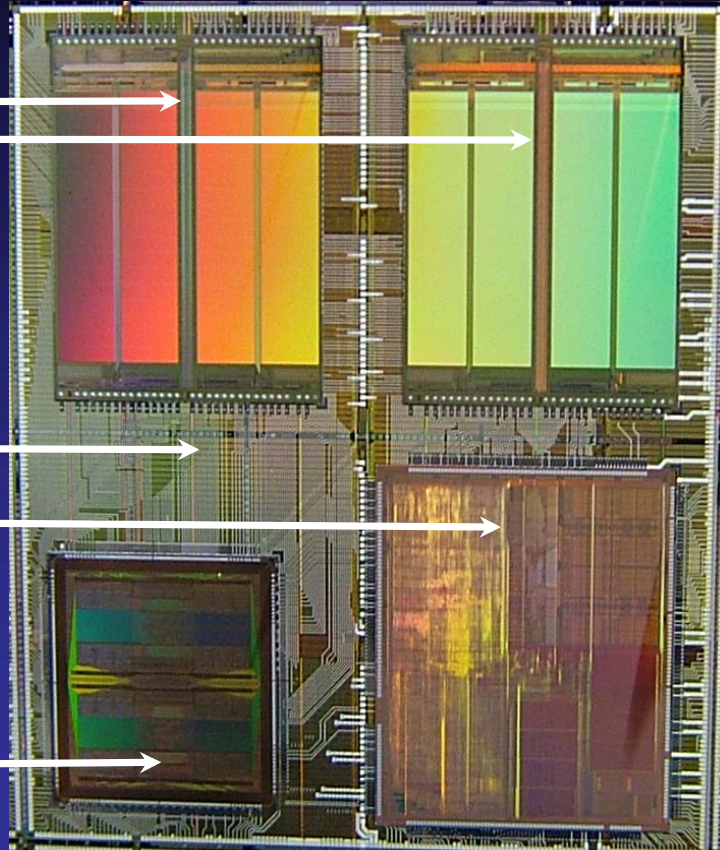
Die 15 x 17 mm

128 Mbit Flash  
x2

1.0 Gbit SDRAM  
(2 x 2 Die Block)

ARM9 ASSP

384 Macrocell  
Reconfigurable  
Logic



Metal  
Interconnect

Memory  
Substrate

Bond  
Interface

Logic Metal

Logic  
Transistors

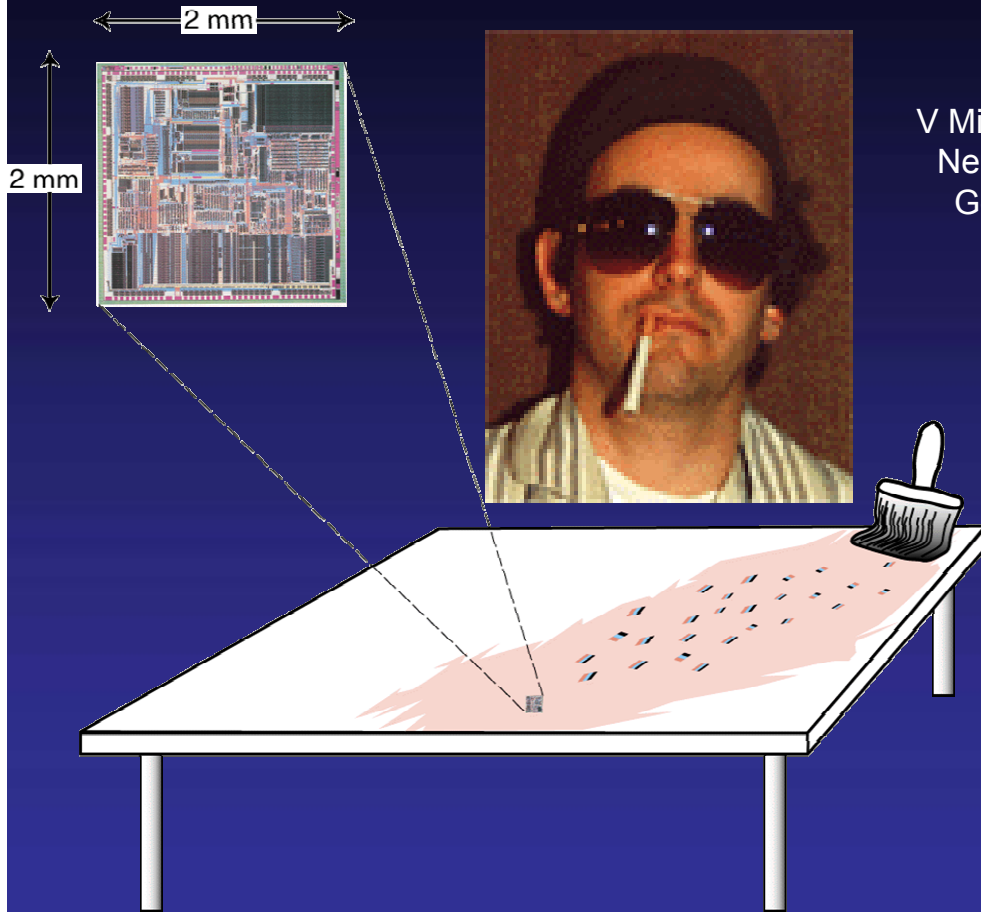
Logic  
Substrate

Ziptronix

# ARM

THE ARCHITECTURE FOR THE DIGITAL WORLD™

# Towards a Billion Processors



Bill Butera  
V Michael Bove Jr  
Neil Gershenfeld  
Gerry Sussman  
Ted Adelson  
MIT

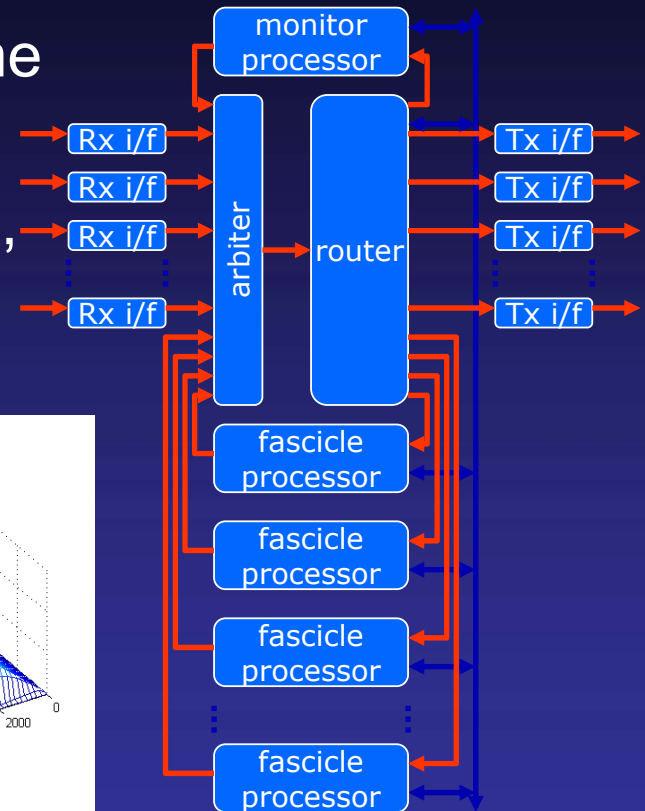
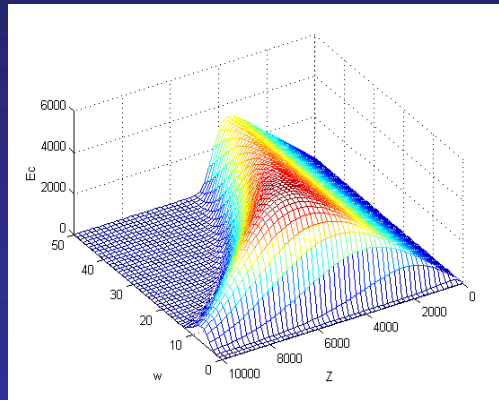
- Smart Motes  
or
- Micro GRID  
computing

**ARM**

THE ARCHITECTURE FOR THE DIGITAL WORLD™

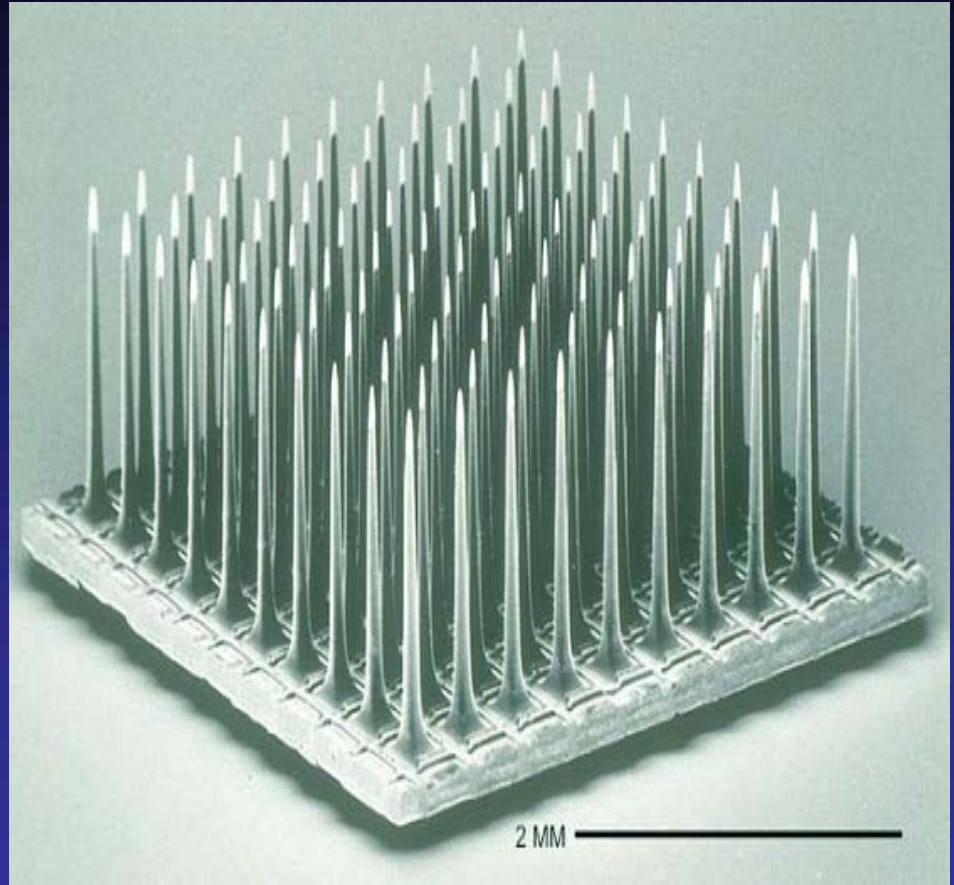
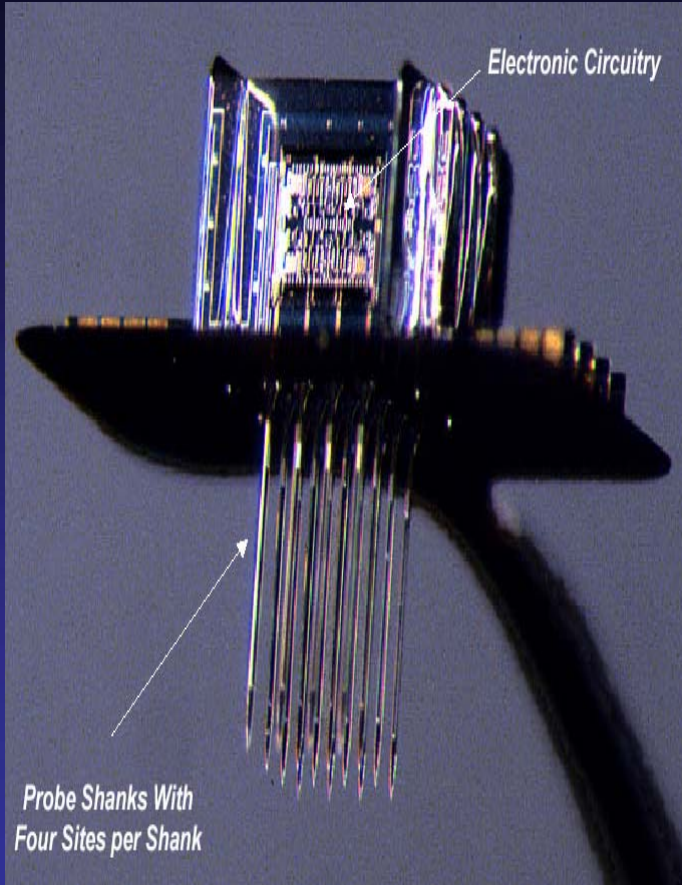
# Neural Systems Engineering

- Long-term goal:
  - Model 1 billion neurons in real time
- Short-term goals:
  - Understand how to build complex, asynchronous neural systems
- Point-to-point links
  - Delay-insensitive
  - Robust
  - Low power
- Async packets
  - Source ID routing
  - ID is payload



Steve Furber Manchester University

# Last Year ..... This Year



Cyberkinetics inc

ARM

THE ARCHITECTURE FOR THE DIGITAL WORLD™