Computer Architecture: The Beginning

Jaynarayan Tudu

Computer Science and Engineering
Indian Institue of Technology Tirupati, India

CS5202 - Lecture 0
18^{th} Jan, 2021

VLSI Past and Lesson Learnt

- It all started with Moore's Law
- The corollary of Moore's Law
  - Performance
  - Power, Energy and Temperature (the Dennard Scaling)
  - Reliability
  - Complexity of the Design
- Evolution in computing

CMOS Transistor

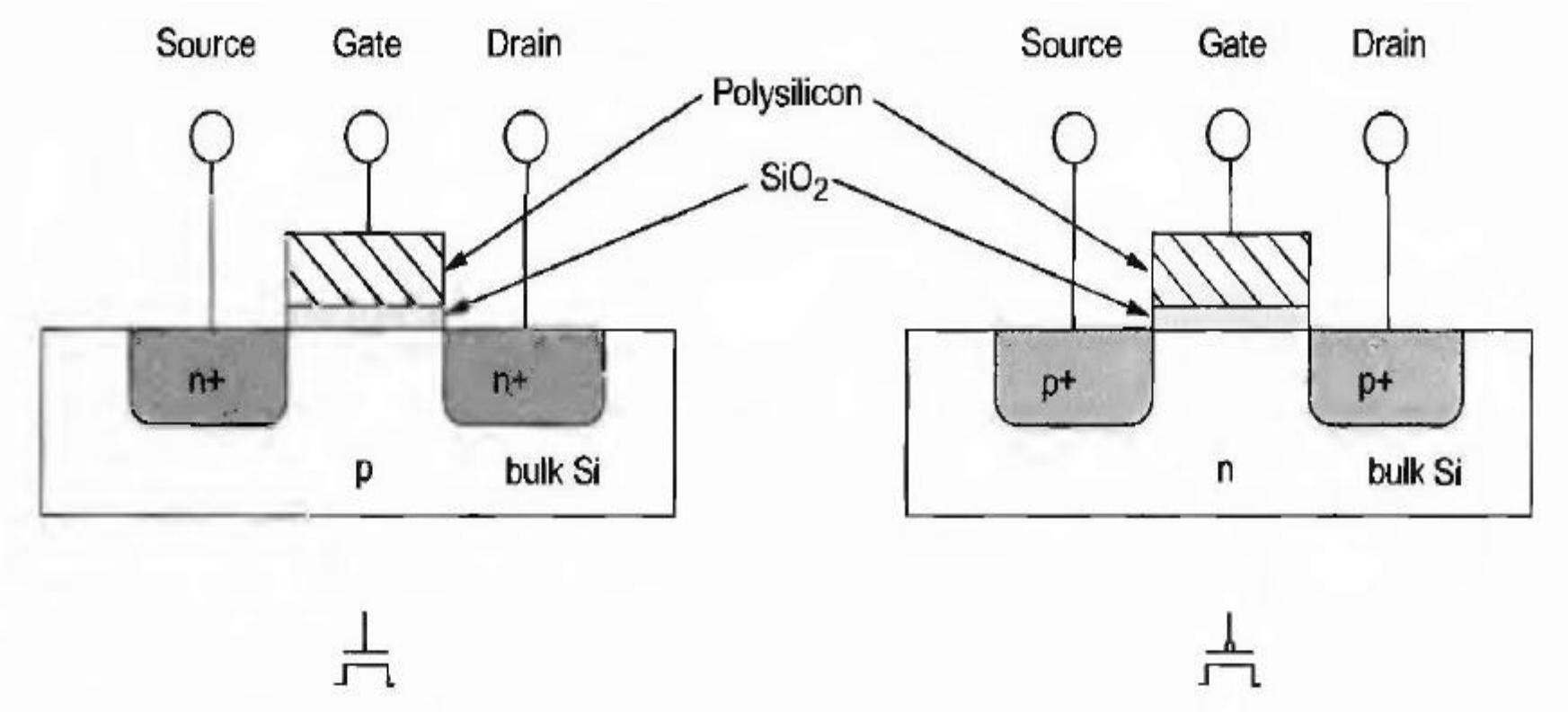


Figure : nmos and pmos transistors [1]

[1] Taken from: CMOS VLSI Des, Weste and Hariss

Transistor and IC History

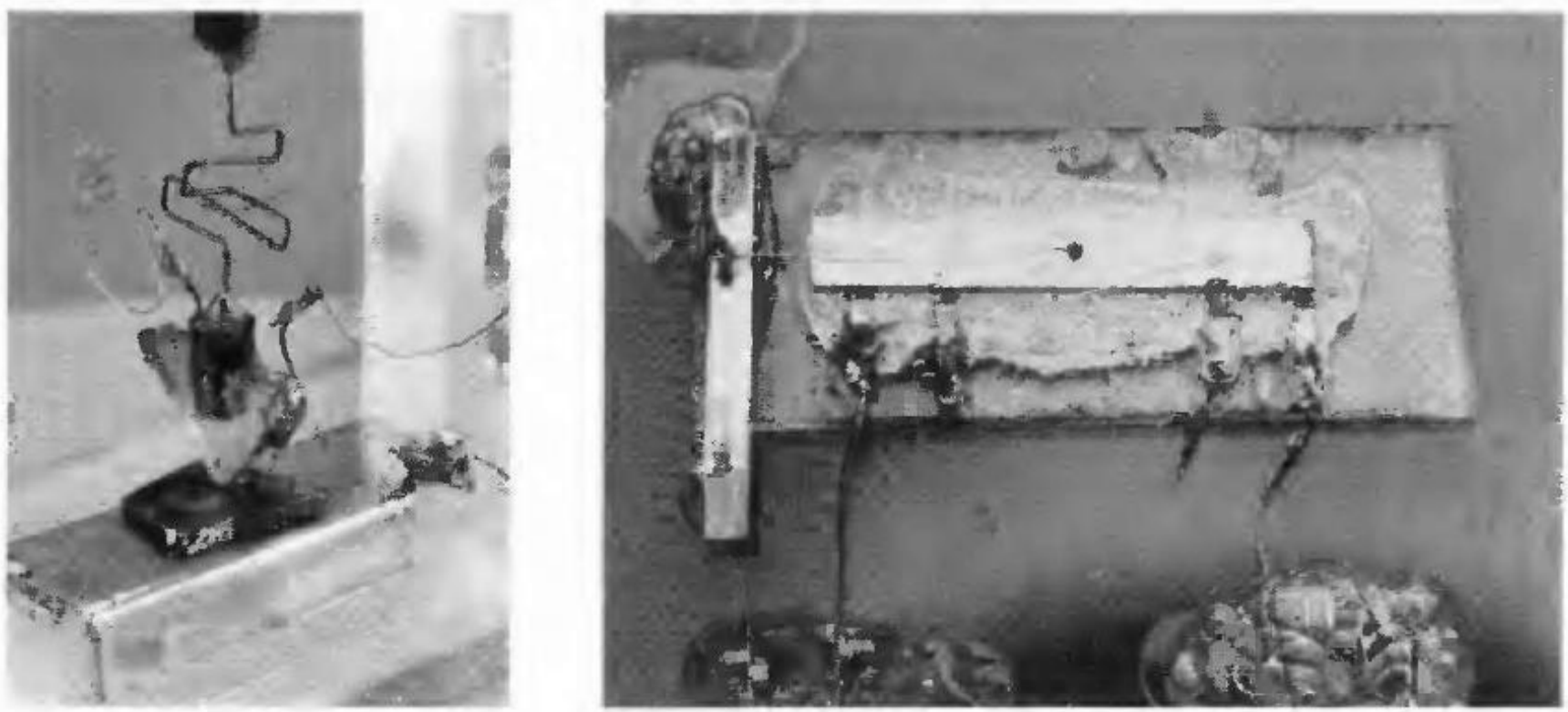
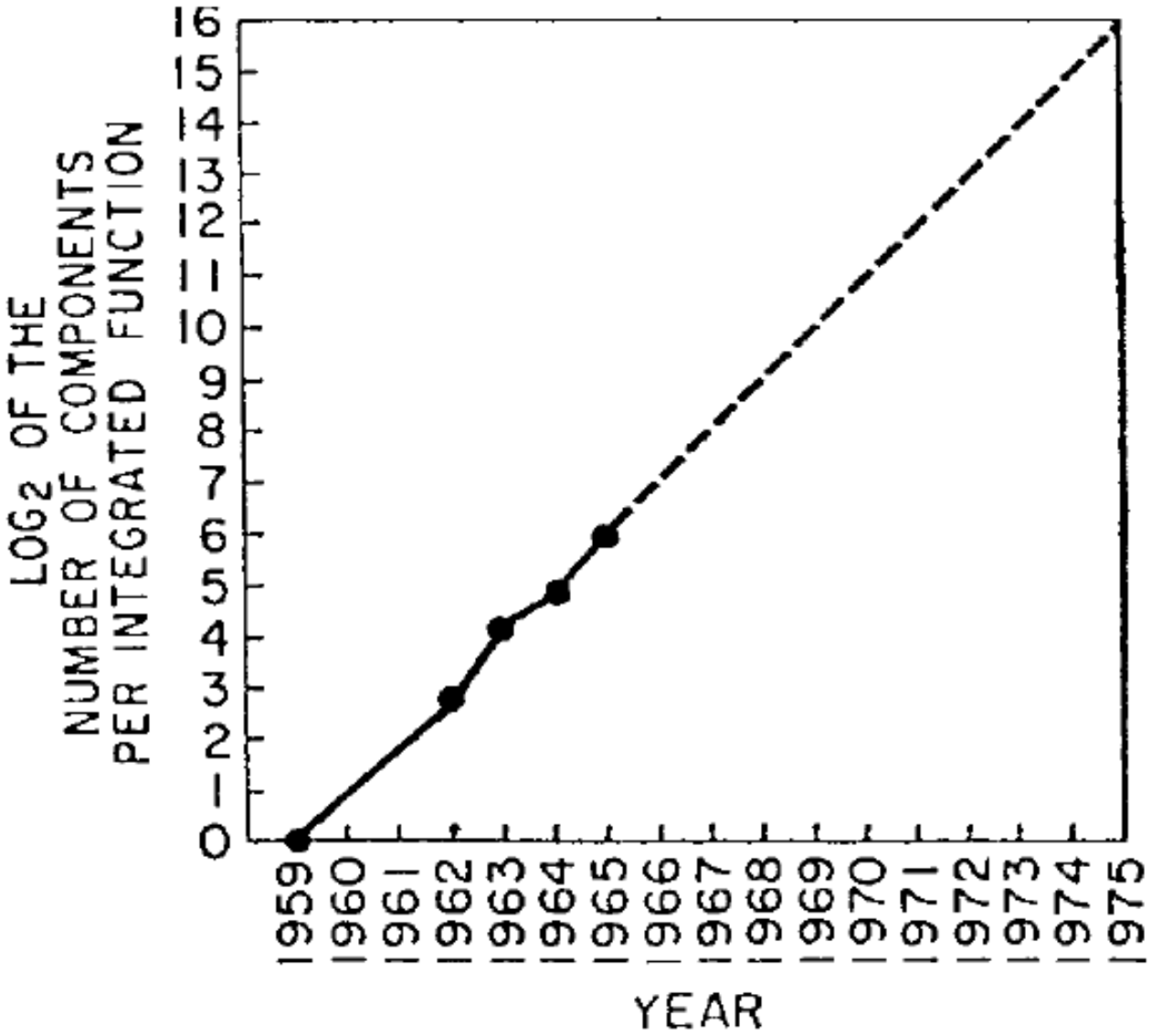


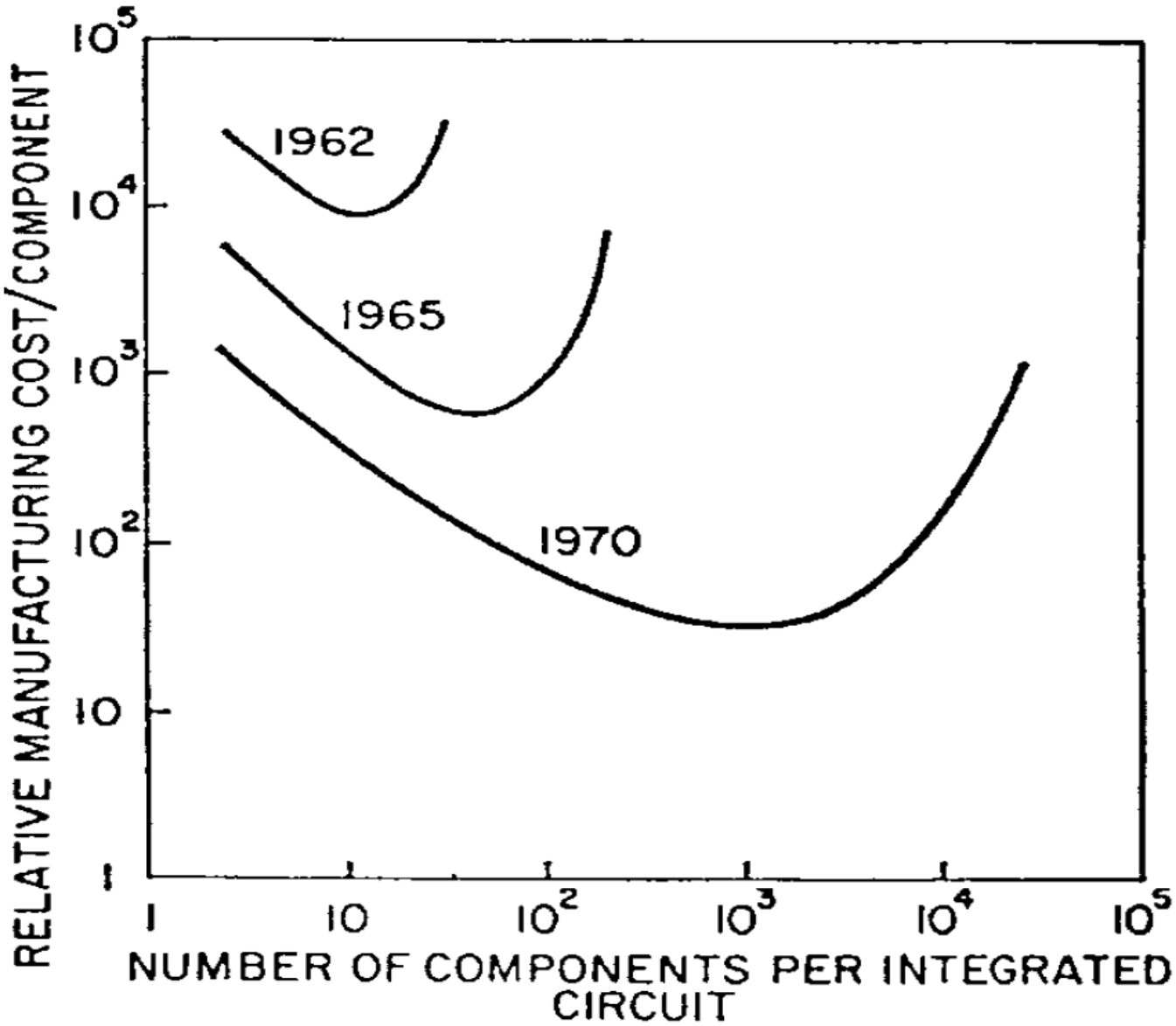
Figure : The first transistor and IC built at AT&T Bell Lab [2]

[2] Taken from: CMOS VLSI Des, Weste and Hariss

Transistor Scaling: Moore's Law

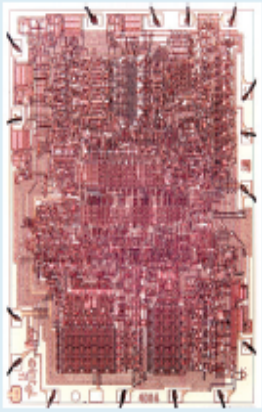


Manufacturing Cost: Moore's Law

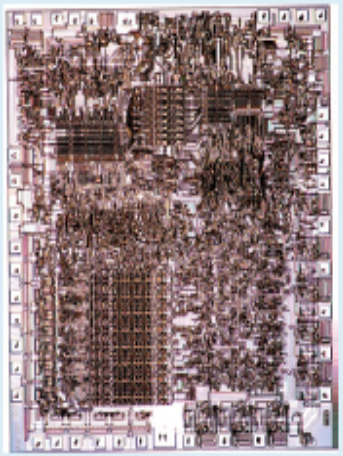


Transistor count

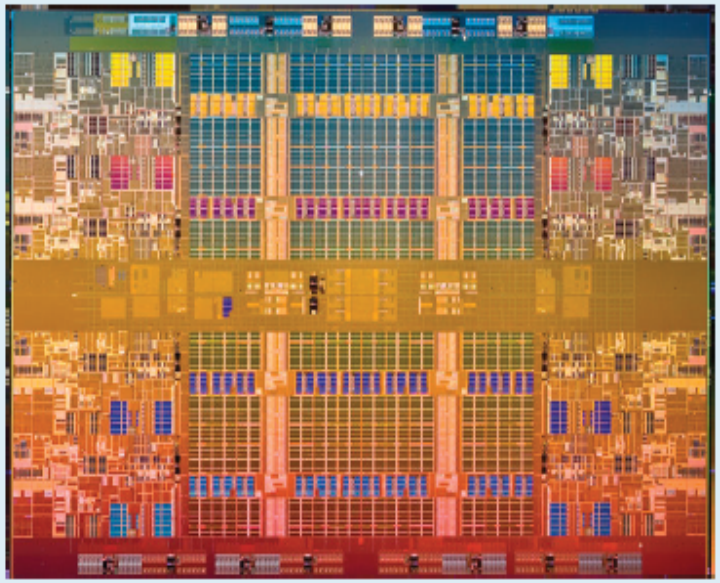
Figure 1. Evolution of Intel microprocessors 1971–2009.



Intel 4004, 1971
1 core, no cache
23K transistors



Intel 8088, 1978
1 core, no cache
29K transistors



Intel Mehalem-EX, 2009
8 cores, 24MB cache
2.3B transistors

Figure : Its is now in billion era

Transistor Production

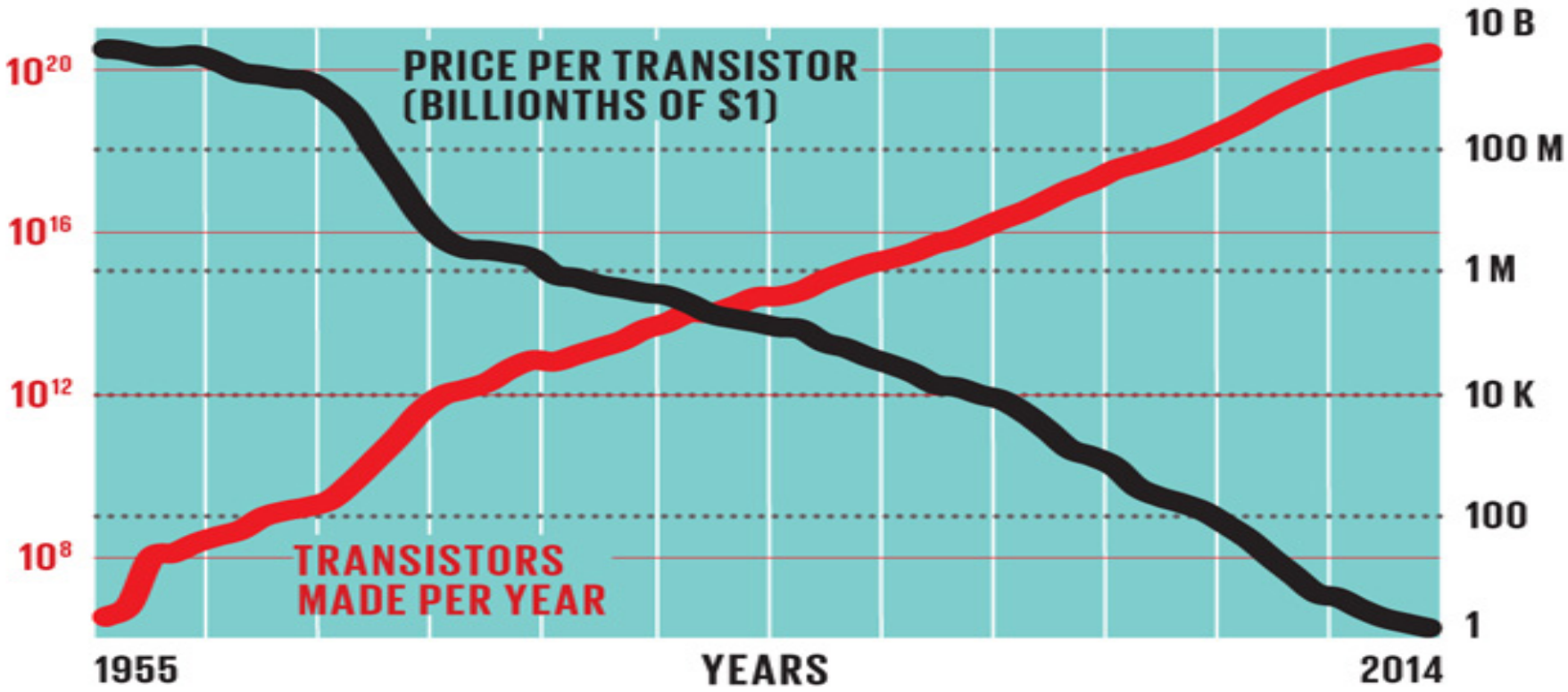


Figure : It is growing every year [IEEE Spectrum]

Transistor Technology Trend

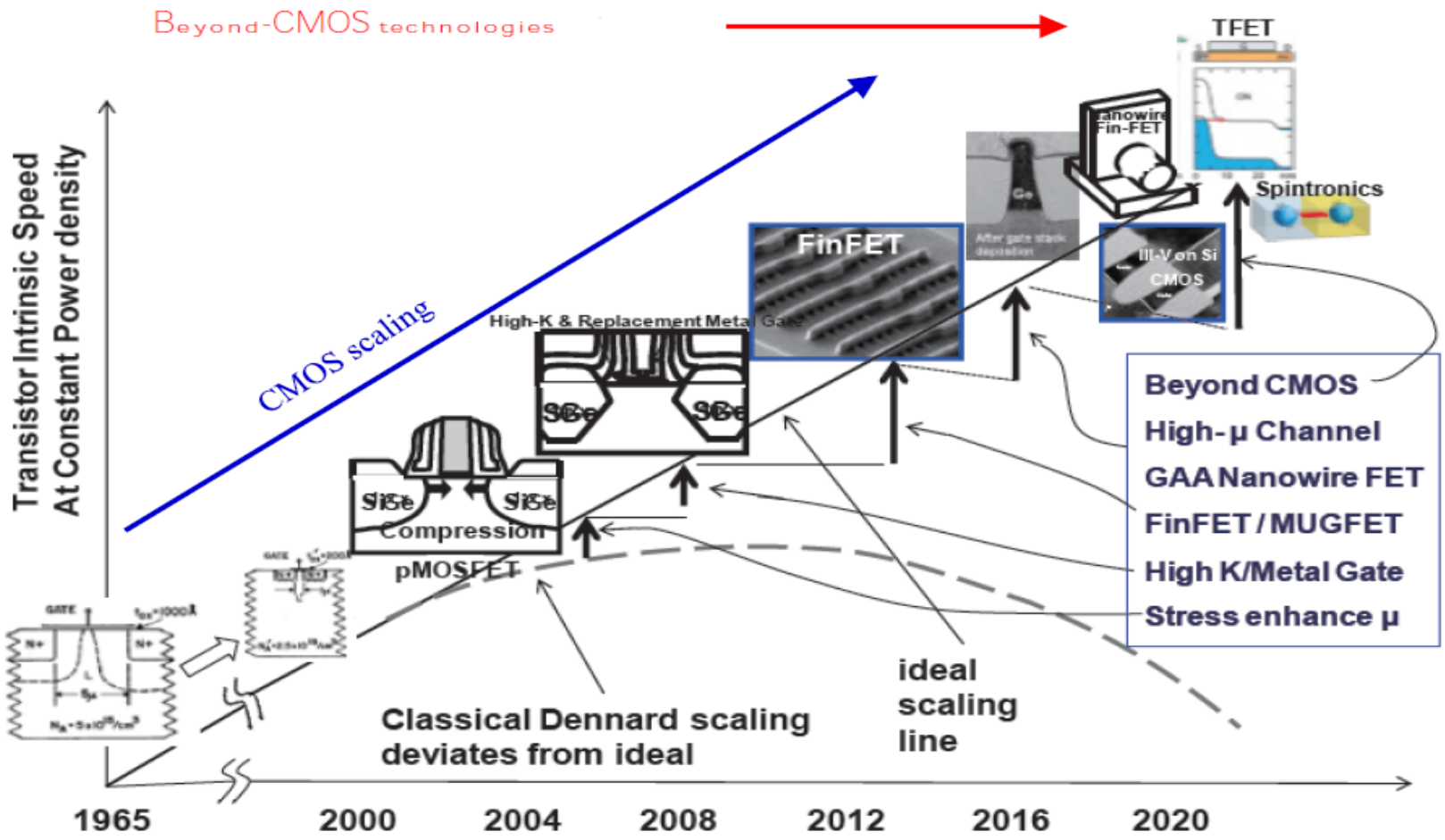


Figure : Scaling driven technology

Trends in Functionality in Microprocessor

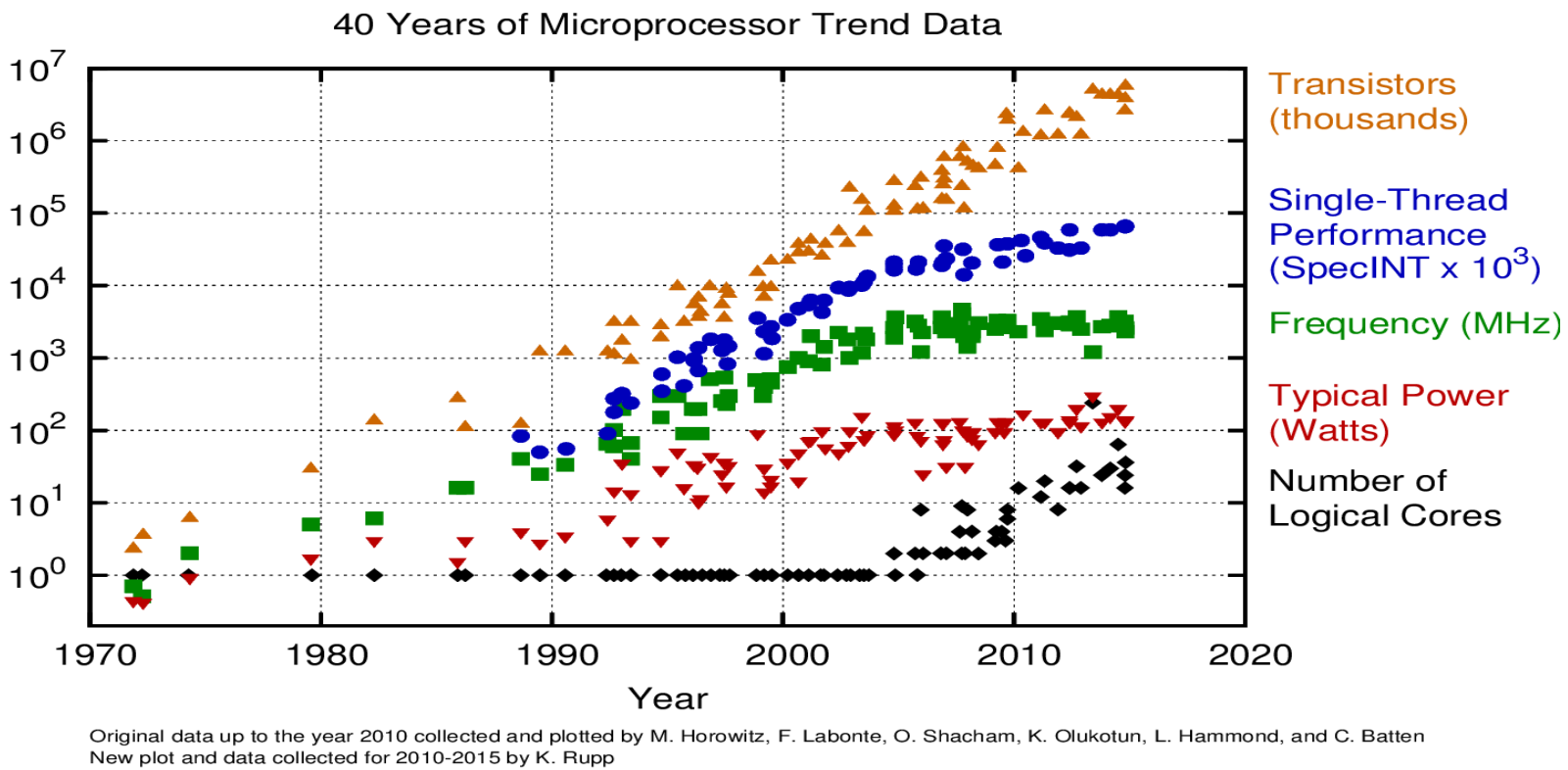


Figure : Microprocessor Functionality

The highest till date: IBM P6 at 5.0 GHz.

Frequency [ITRS 2013]

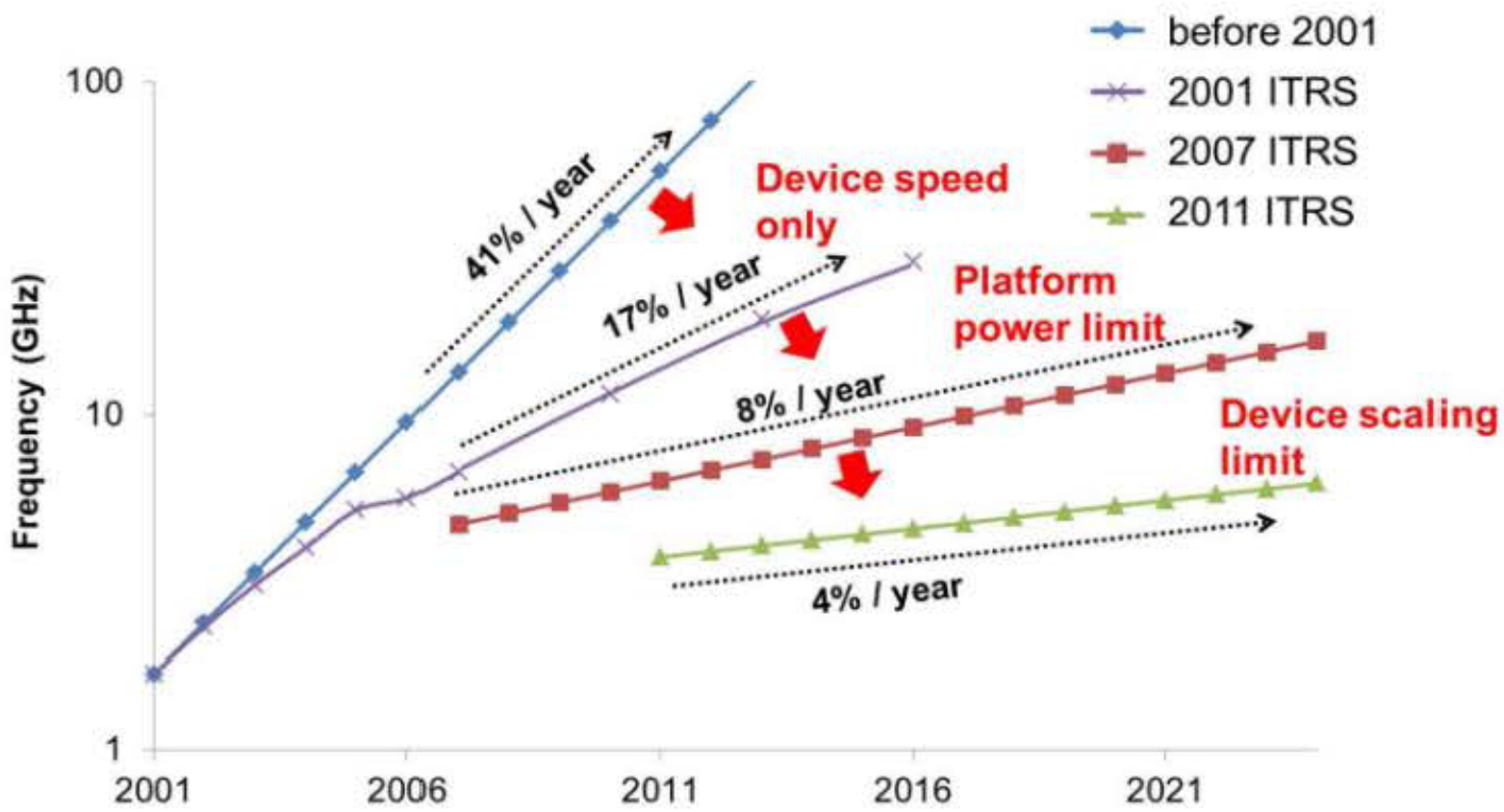


Figure : the rate has slowed down

Intel Microprocessor Trend: Clock frequency

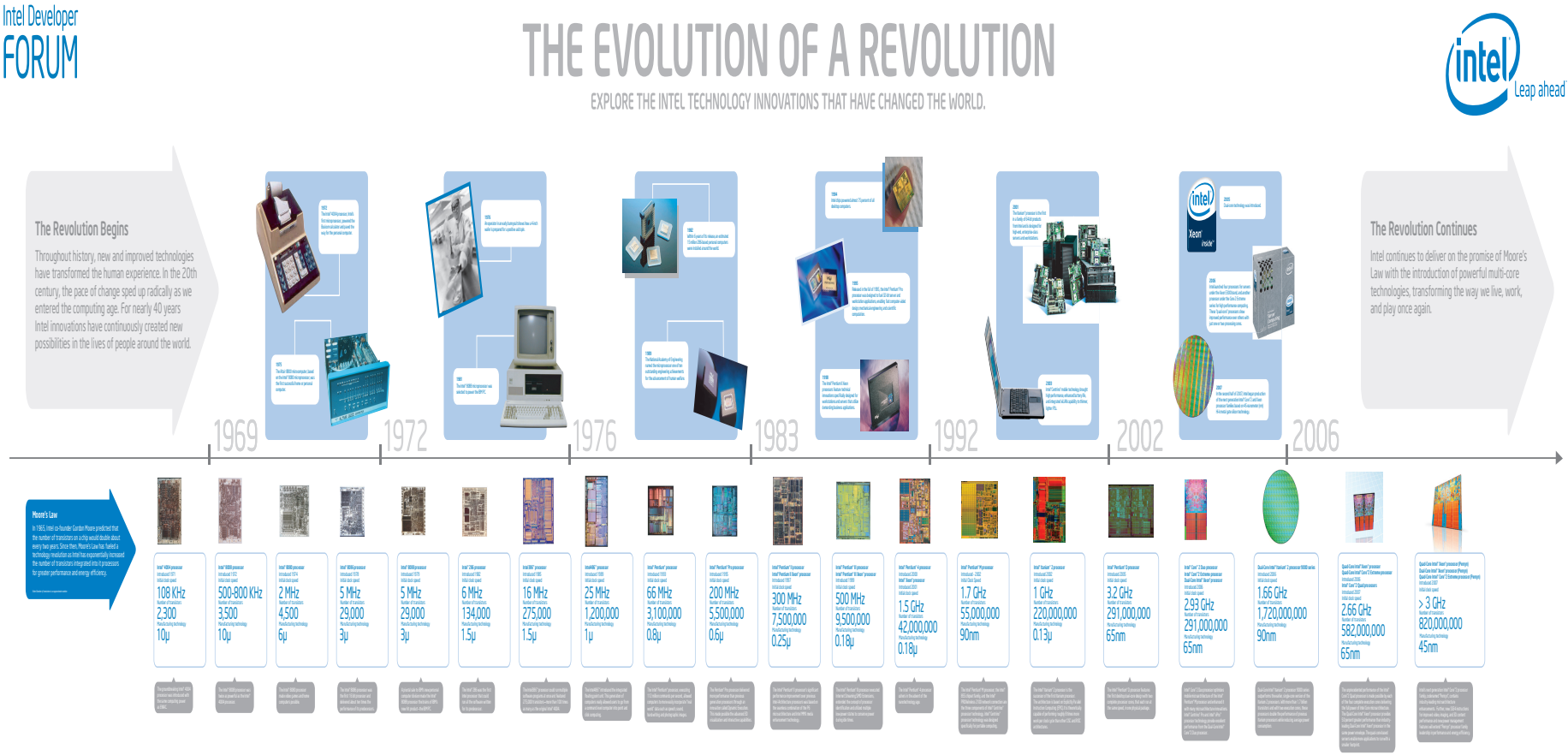


Figure : Clock frequency has been consistently in rise!

Power Density Trend [ITRS 2015 — A Kahng, UCSD]

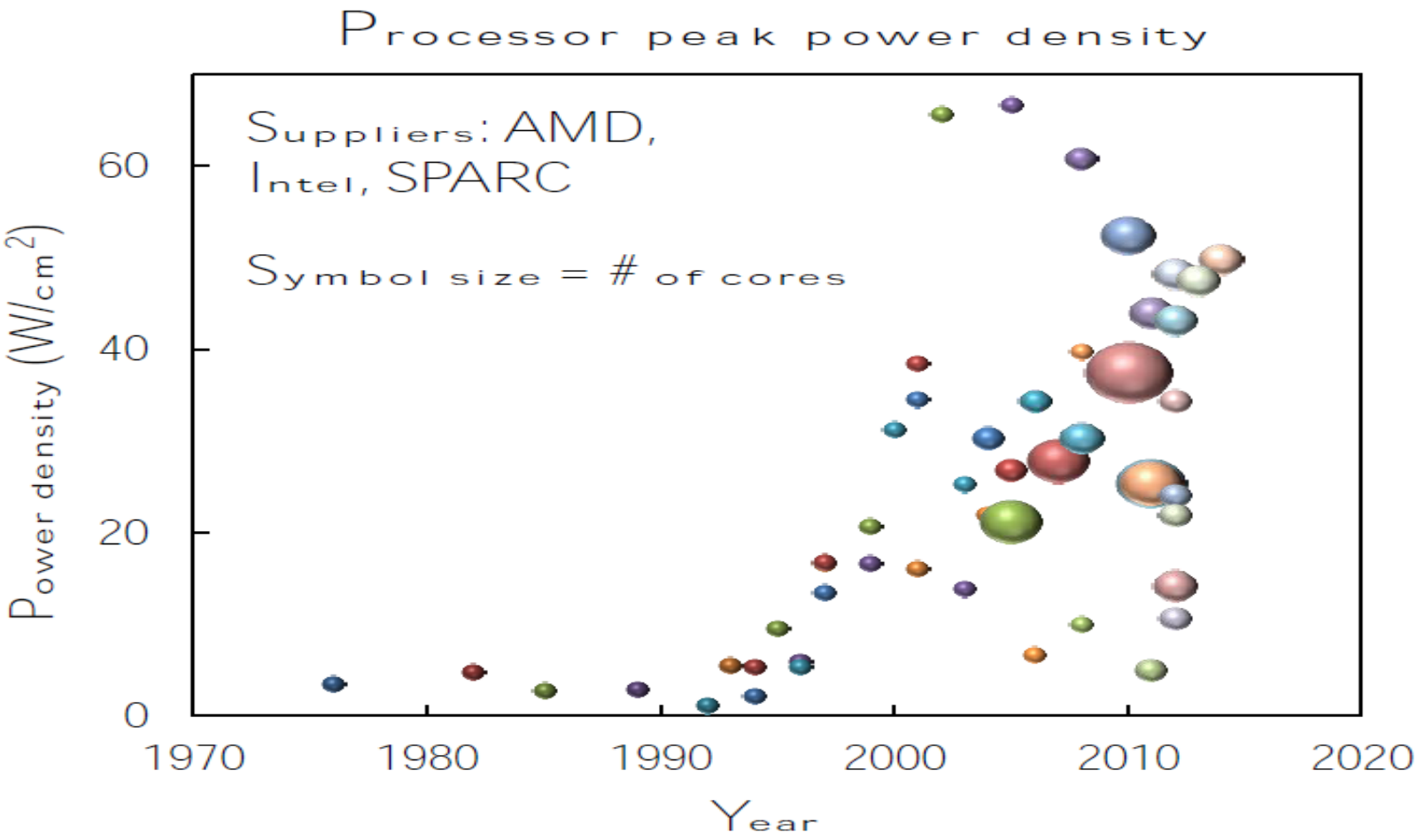


Figure : Peak power density trend

Power Consumption [Borkar et al]

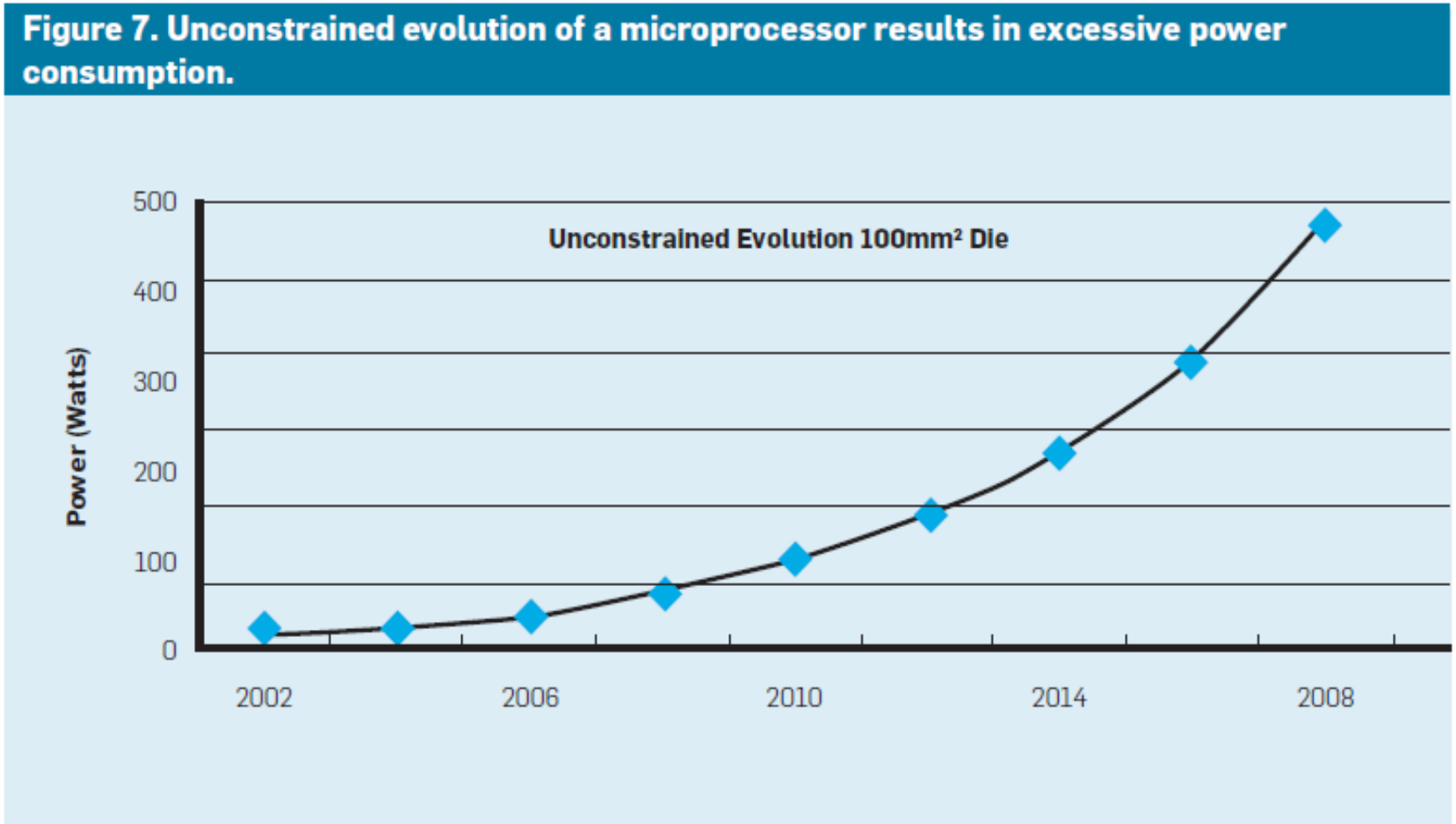


Figure : Processor power consumption

Heat Flux [IBM — ITRS 2015 meeting]

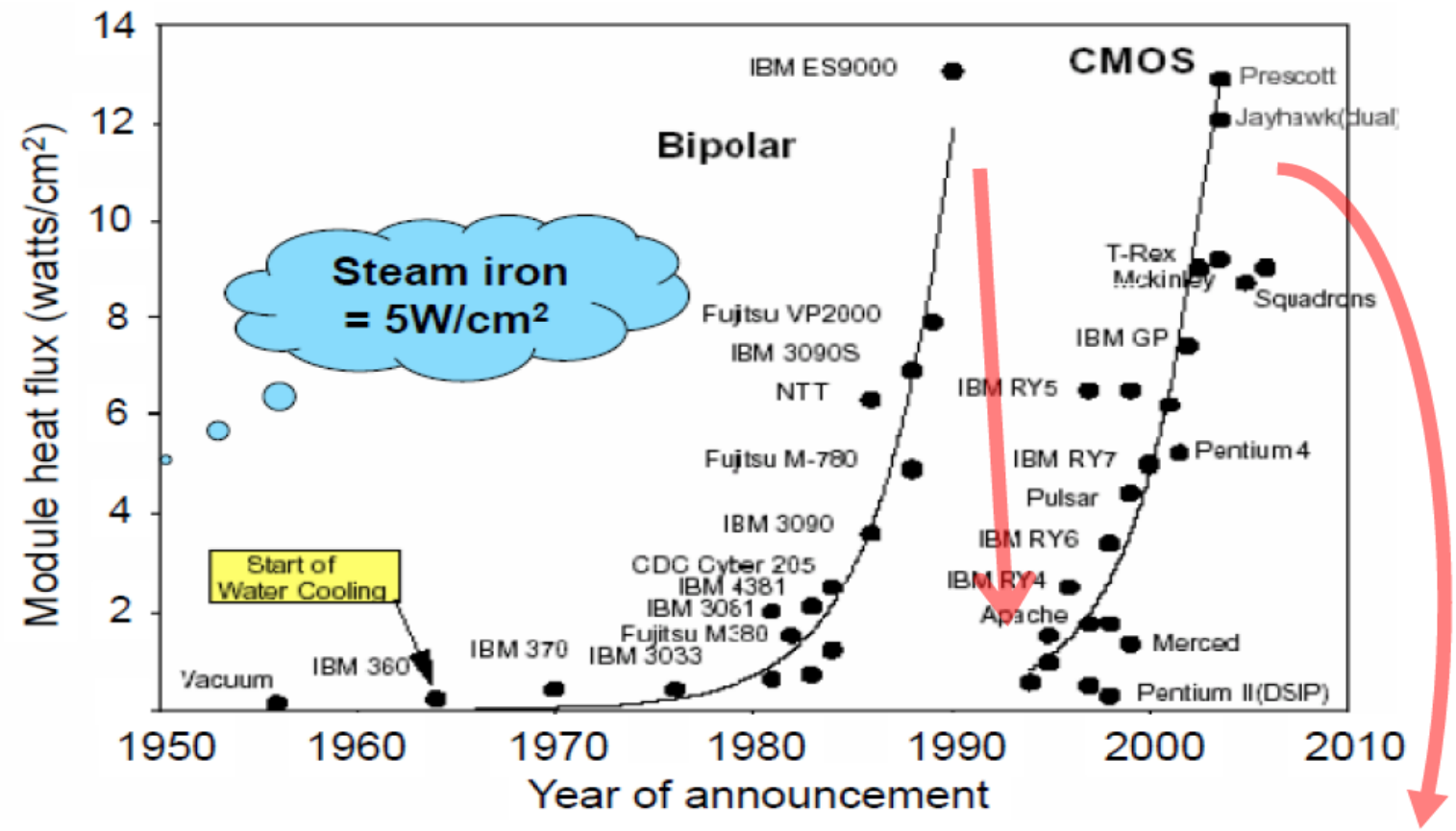
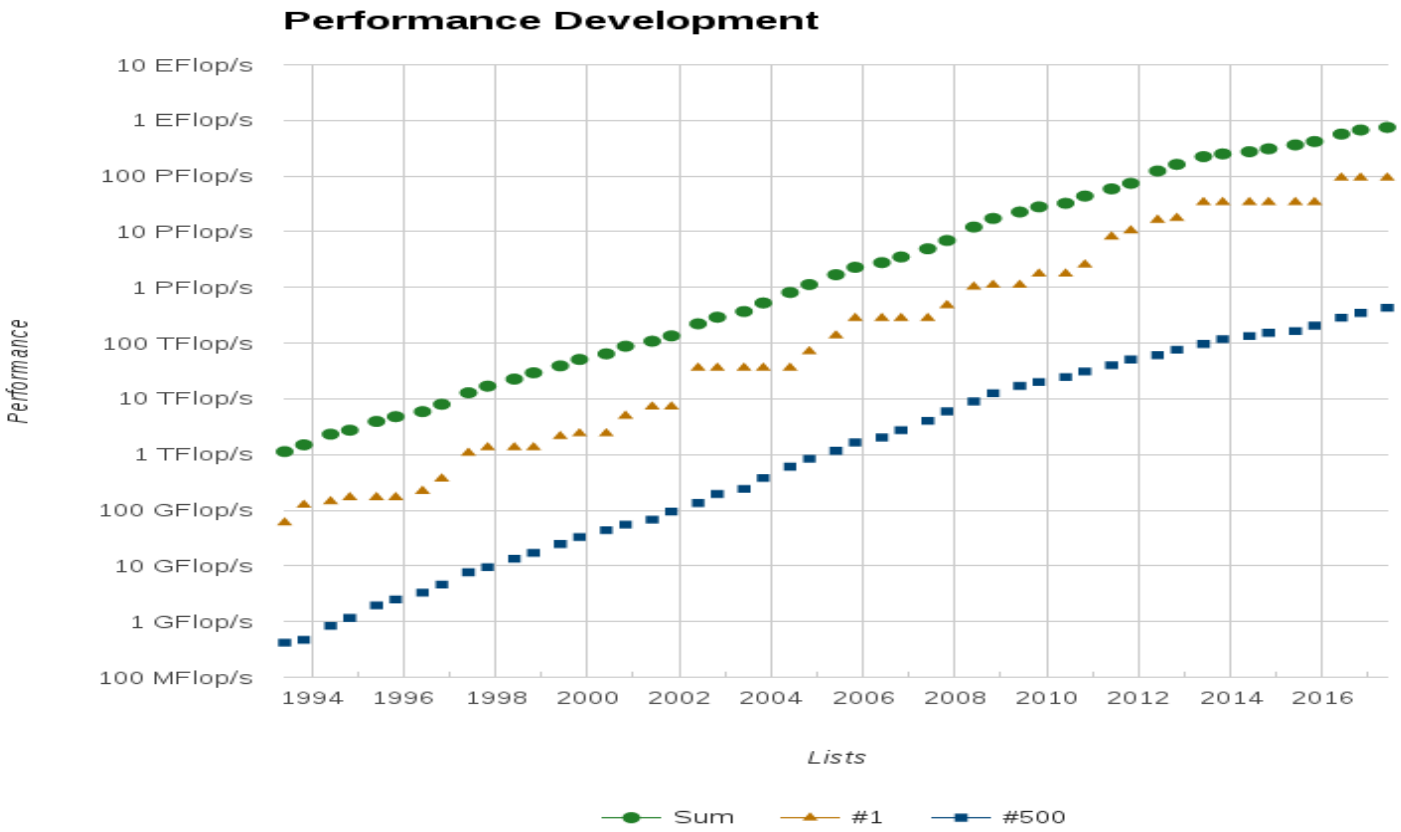


Figure : Heat flux Watt/area

There is a need of new device technology to tackle the heat flux!

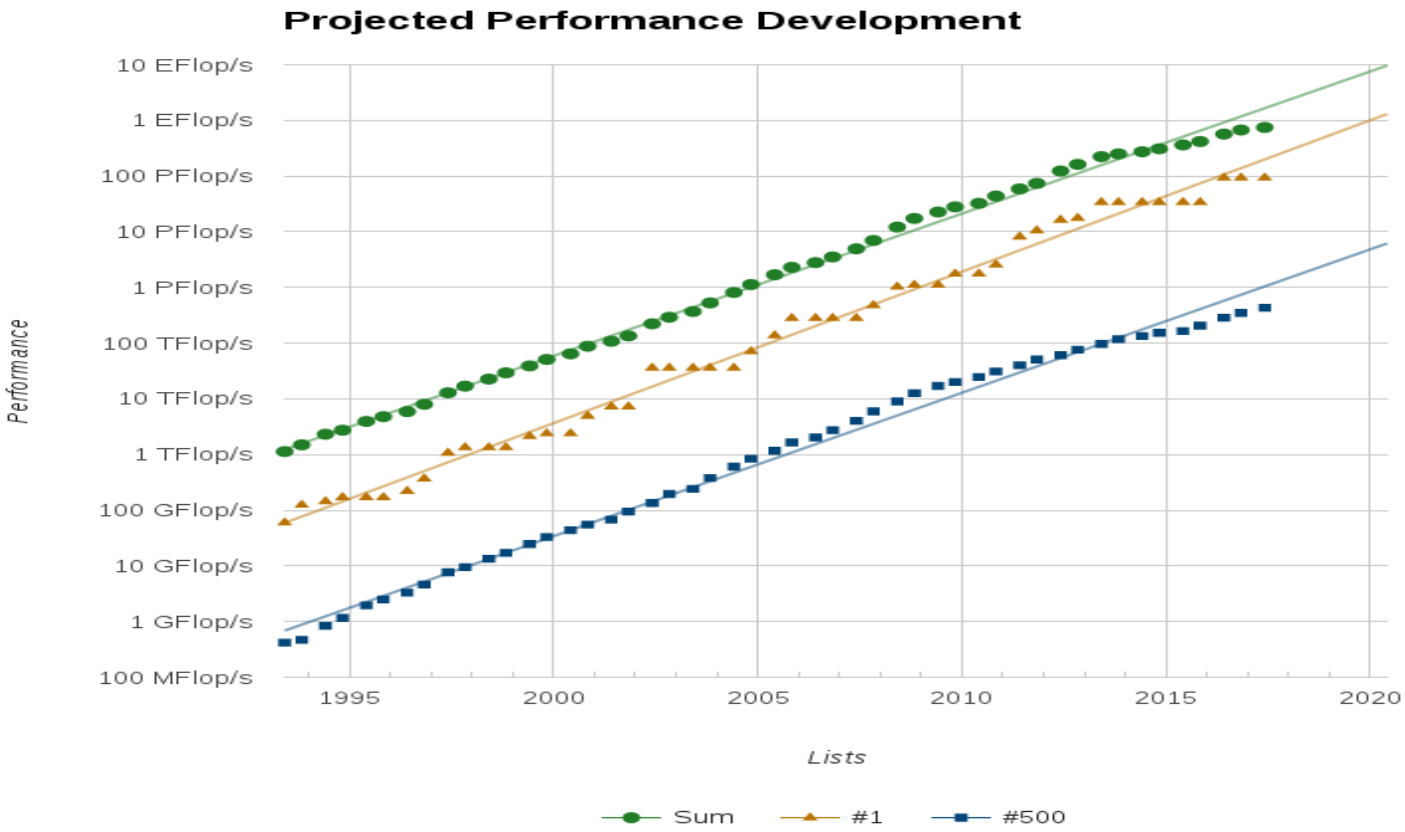
Exascale Computing: Trends in Top500.org

- Power and Energy will pose a major challenge [Borkar 2011].
- GWatt consumption is expected.

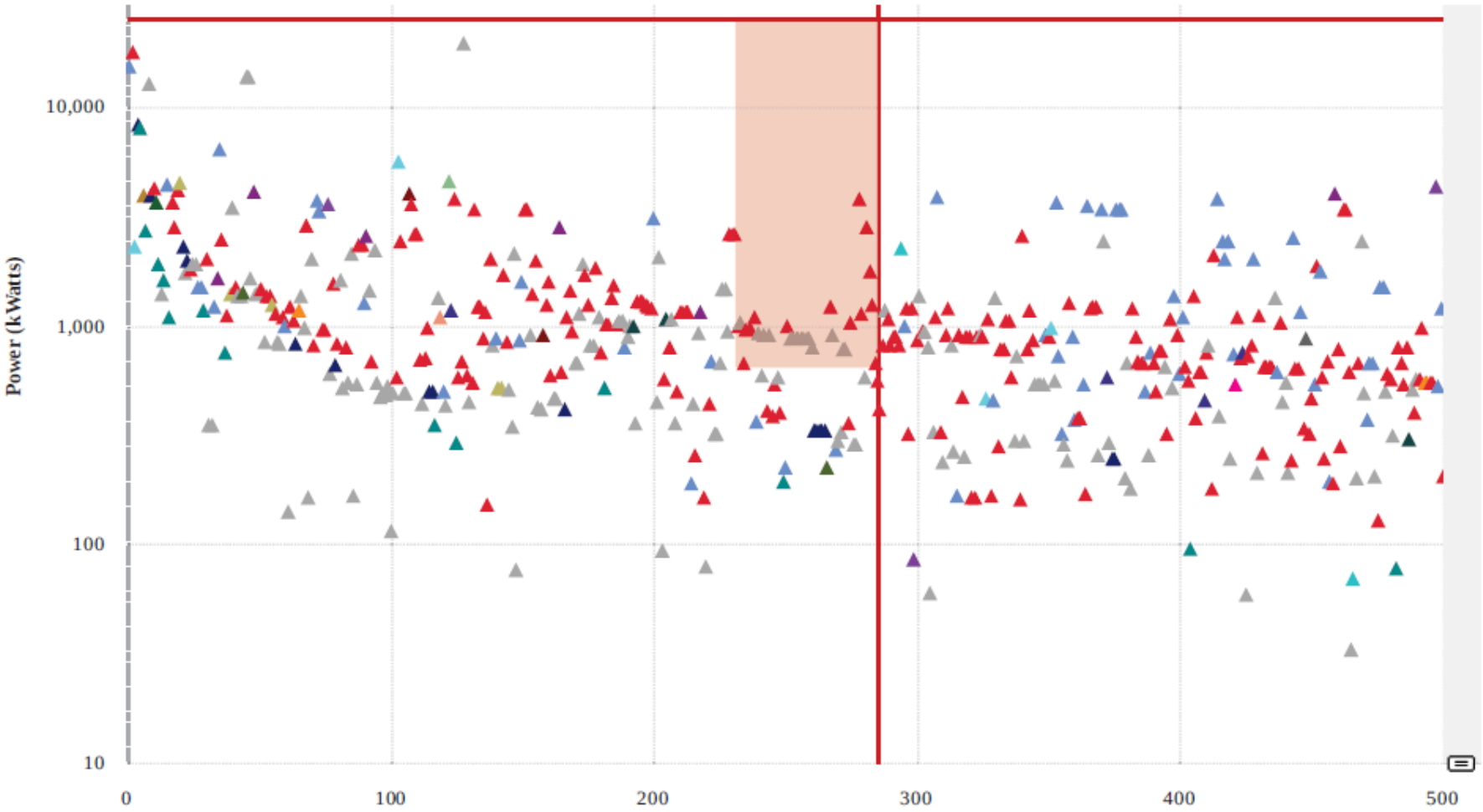


Exascale Computing: Trends in Top500.org

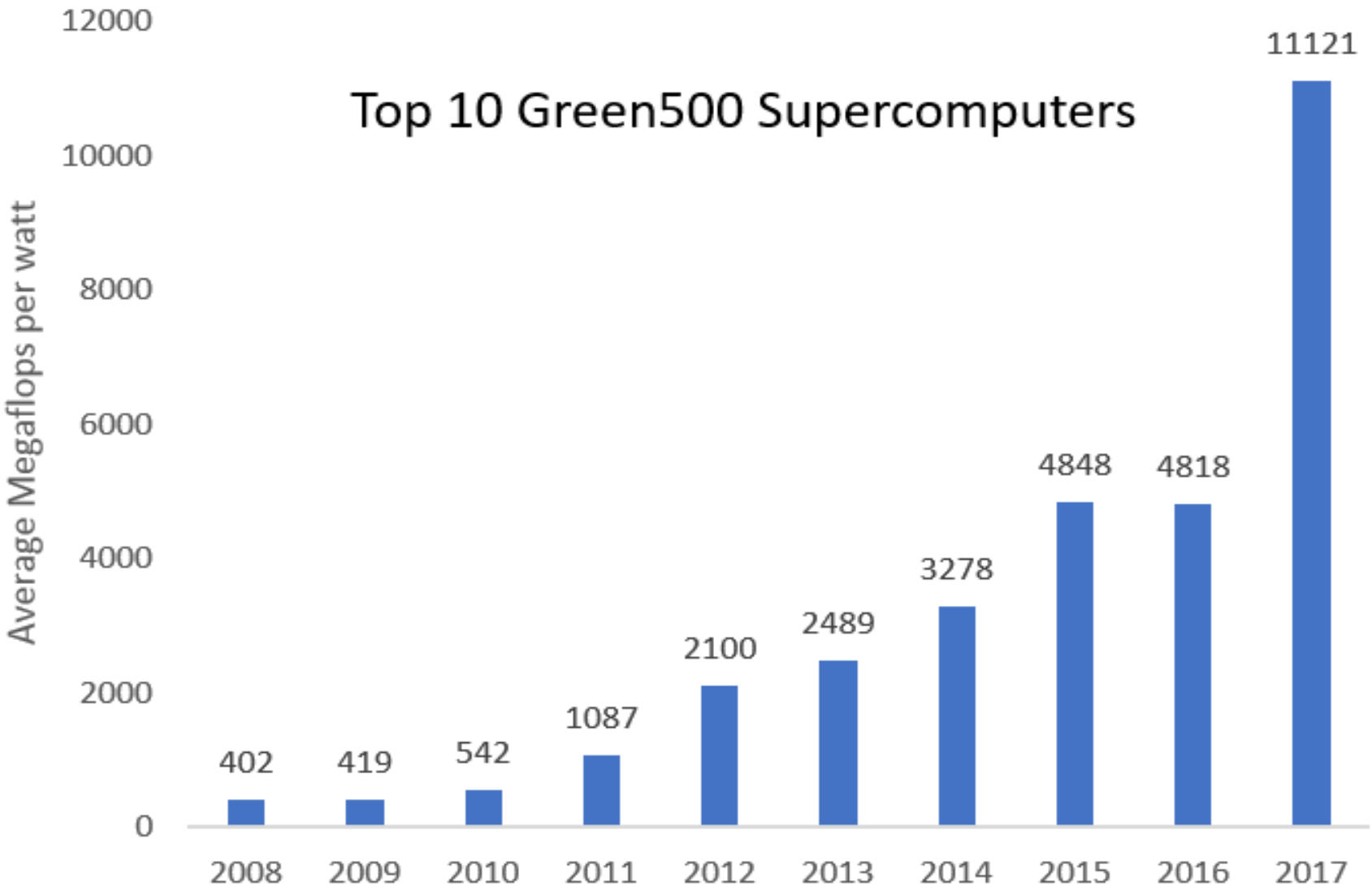
- Power and Energy will pose a major challenge [Borkar 2011].
- GWatt consumption is expected.



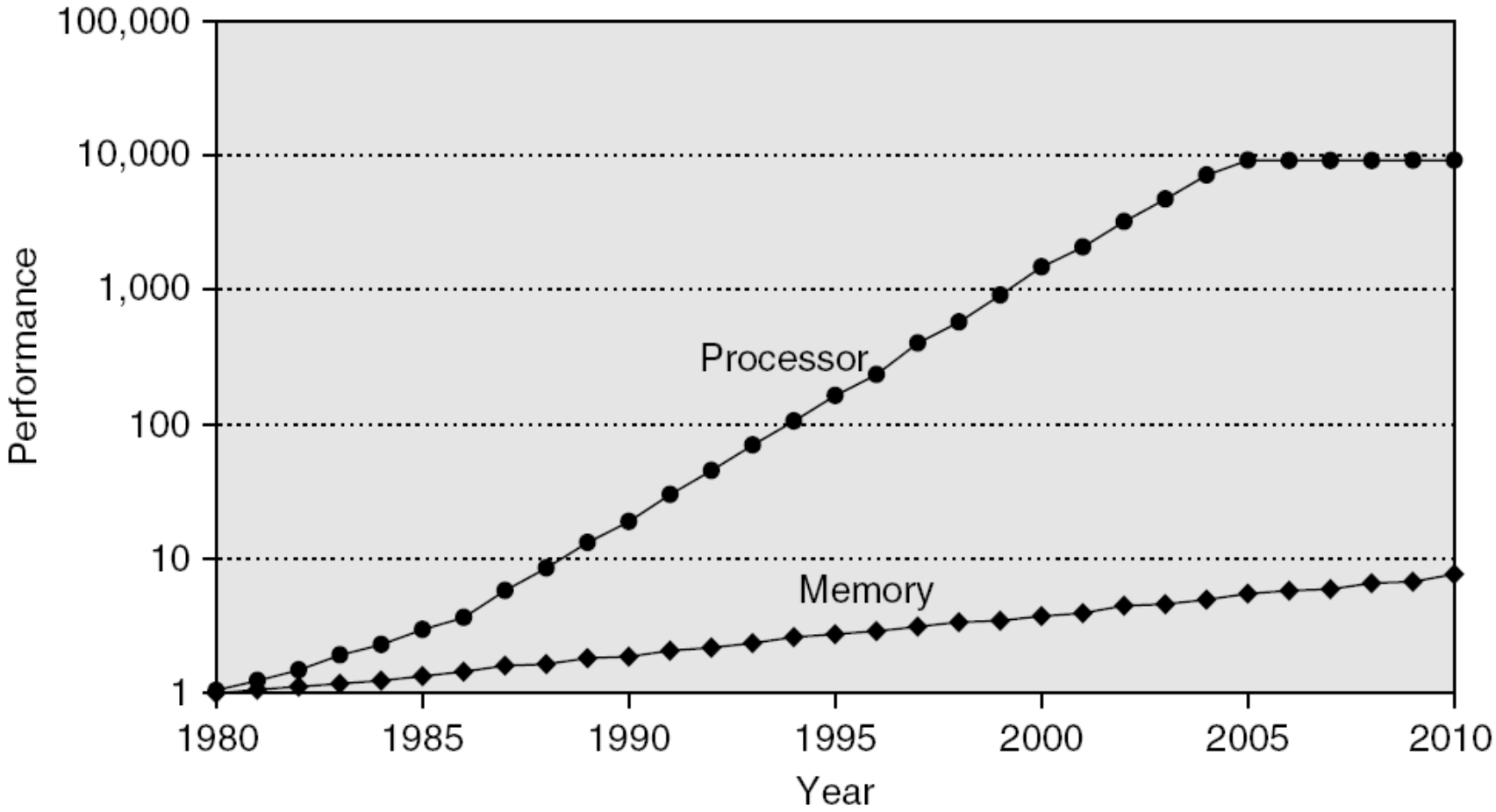
Power Trend in Top500



Performance per Watt Trend [top500.org]



Memory Wall: Hindrance in performance



Emerging Memory Technology [ITRS 2015 meeting]

Solution: The idea is to bring in hierarchy system!

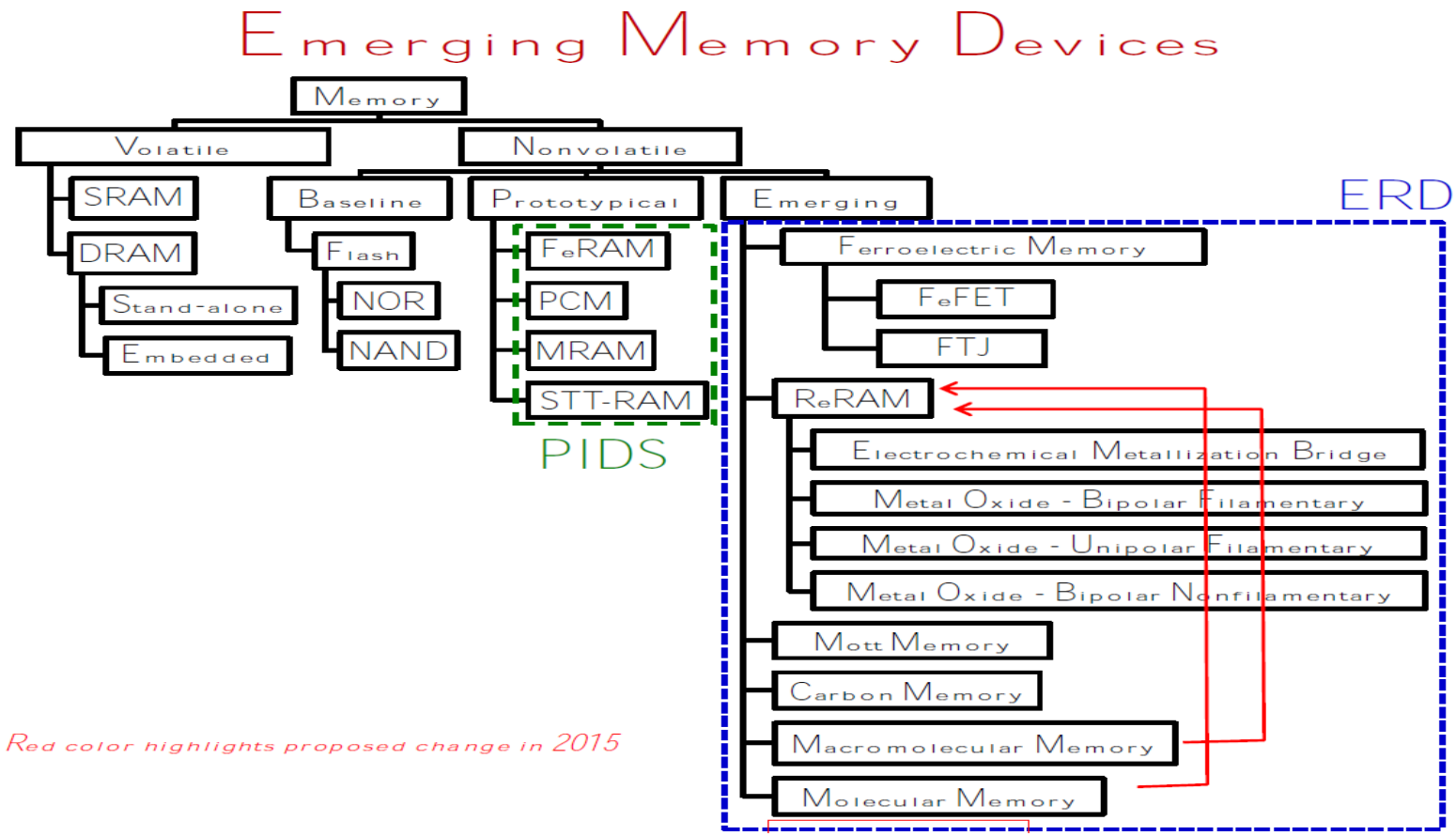


Figure : Focus is on non-volatile memory

On die Cache [Borkar et al]

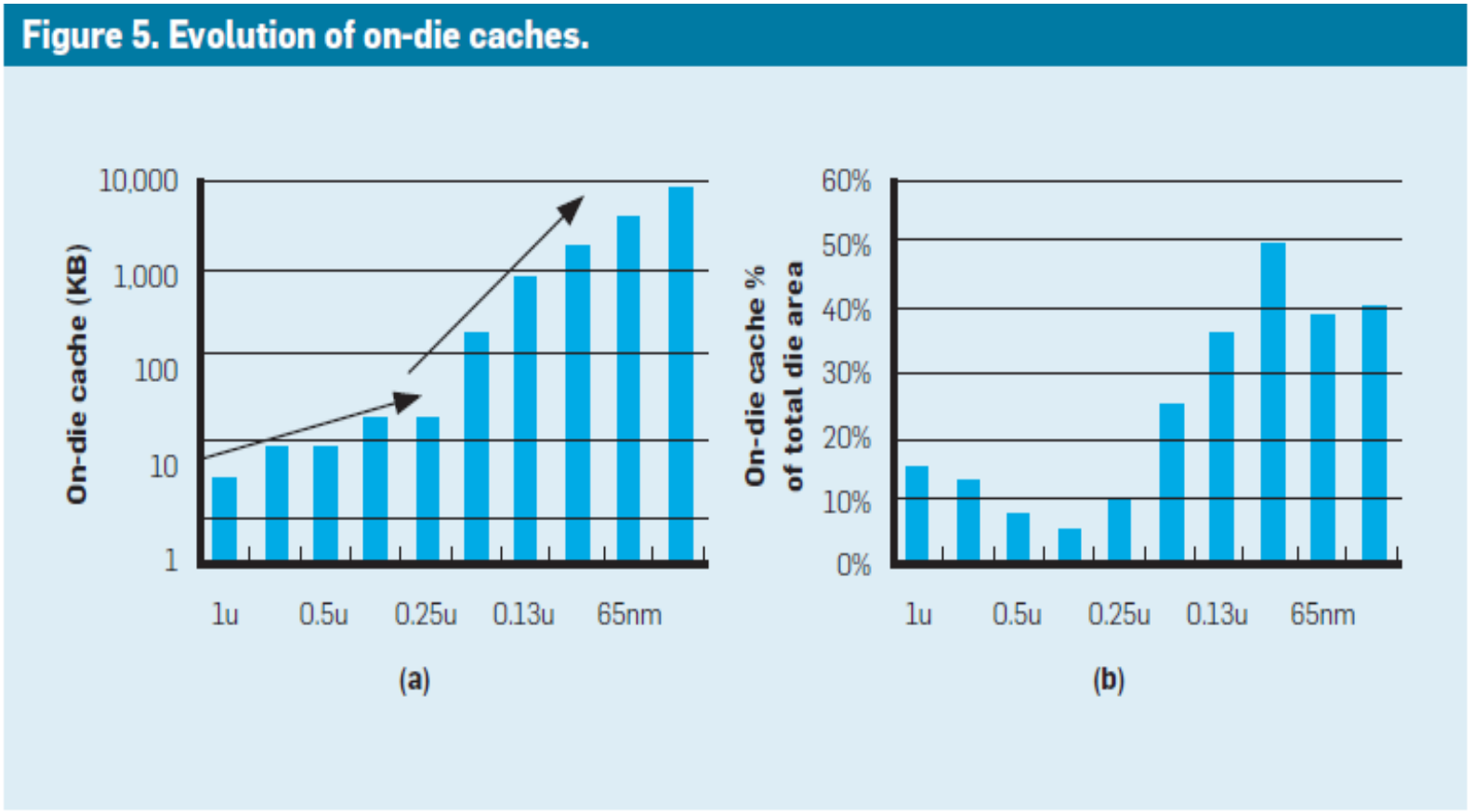


Figure : Suppose to increase

Emerging Architecture

von Neumann

- The traditional microprocessor

Non von Neumann

- Cellular Automata
- Co-located memory-logic [procesor-in-memory, memory-in-logic, nonvolatile logic]
- Reconfigurable computing
- Cognitive computing [neuro morphic, machine learning]
- Statistical and stochastic computing [statistical inference, approx computing]

Emerging Architecture: Data and learning

In Design or test production

- Tensor Processing Unit [Google ML processing for Search and Language]
- Microsoft Azure with Stratix 10 (intel) FPGA
- XPU by Xiling for Machine learning
- IPU (Intelligent Proc Unit) by Graphcore
- DPU (Data flow computnig) by Wavecore
- DLU by Fujitsu

Check out top500 for more information.

Trends in IoT

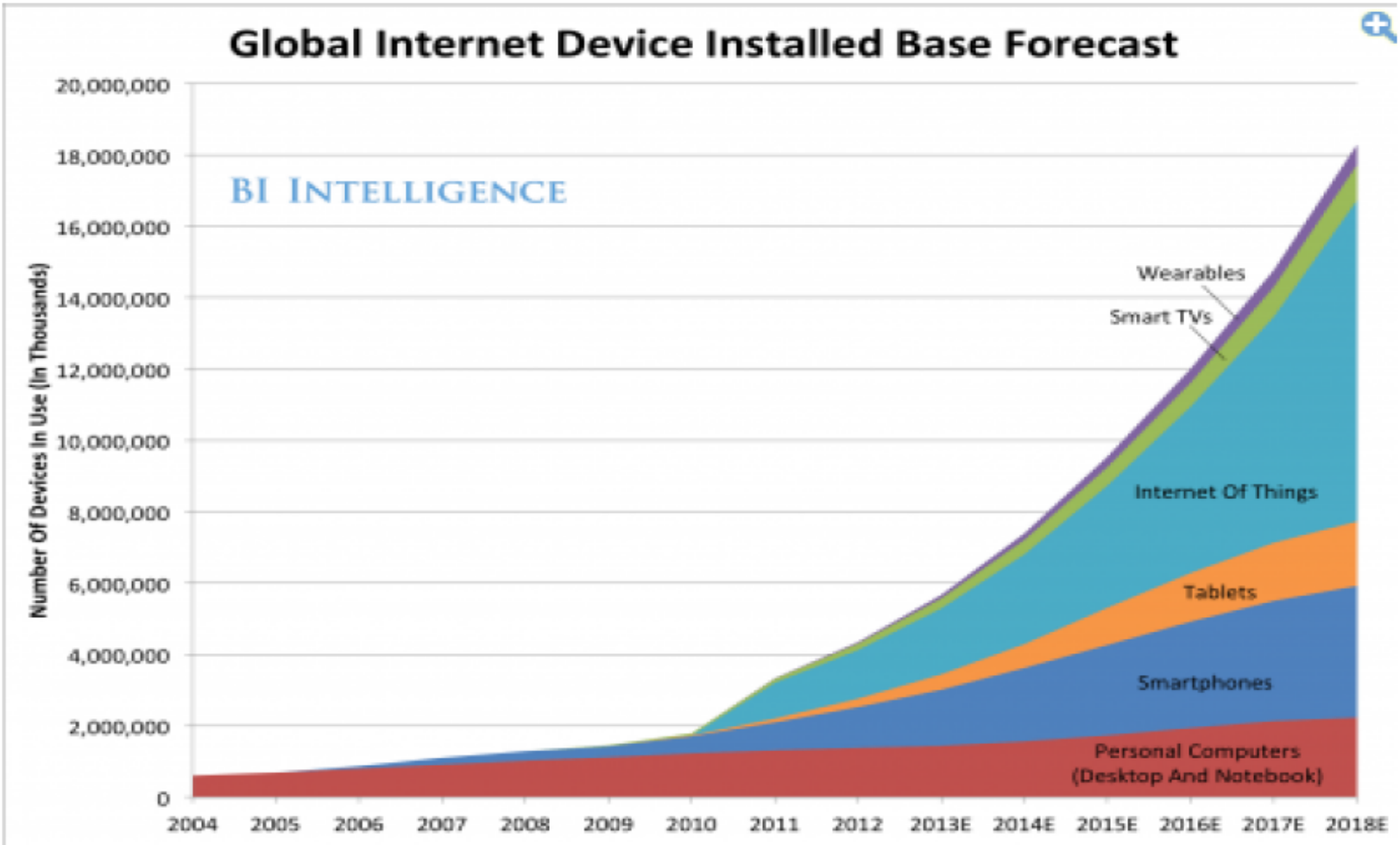


Figure : IoT in use

Summary and Action Items

- There are problems to be solved
- Teaching and research to continue...

The big question

How to push forward the performance while controlling significantly the power/energy dissipation/consumption?

CS5202: Course Outline

- Performance Evaluation [in depth]
- Instruction set Architecture: CISC vs. RISC [in depth]
- Memory system architecture: Main memory and cache architecture [in depth]
- Pipelining and Superscalar Architecture [in depth]
- VLIW and Multi-scalar Architectures [introductory]
- Simultaneous Multi-threaded (SMT) Architecture [in depth]
- Multi-core Architecture [introductory]
- SIMD, GPU Architecture and Accelerator Design [introductory]

Course Schedule and Meeting

- Schedule:
  Slot G
  Monday: 12 - 12:50, Tuesday: 14 - 14:50, Friday: 14- 14:50
- Meeting:
  - TAs: Aditi Palit, PhD student CSE
    cs21d001[at]iittp.******
  - Instructor: Jaynarayan T Tudu
    jtt[at]iittp.*****
    #1, 2nd Floor, Annex building,
    Temp Campus, IIT Tirupati
    http://jayresearch.github.io

Course Evaluation

Quizes (Qz1: 10 and Qz2: 10)	20%
Class Test and Participation	10%
Assignment (Programming/simulator)	10%
Project (Group of 2)	20%
Final Test	40%

Reference and Reading Materials

- J. L. Hennessy and D. A. Patterson, Computer Architectures: A Quantitative Approach, Morgan Kaufmann Publishers, 5^{th} Edition.
- J.P. Shen and M.H. Lipasti, Modern Processor Design, MC Graw Hill, Crowfordsville, 2005
- Current Literature (from ISCA, Micro, HPCA, ICCD, and IEEE Trans. on Computers, IEEE Architecture Letters)

The other resources:

- http://pages.cs.wisc.edu/ arch/www/
- Prof. Onur Mutlu: https://users.ece.cmu.edu/ omutlu/
- Almost all universities have very strong research group
- SPEC Benchmark: https://www.spec.org/benchmarks.html
- Simulators: simplescalar, sniper, gem5, gpgpu-sim, tejas
- DBLP: https://dblp.uni-trier.de/

Research in India

- IISc Bangalore: Prof Matthew Jacob, Prof. R Govind Rajan, Dr. Uday Bandagolu, Dr. Arka Basu, Prof. S K Nandi.
- IIT Bombay: Prof. Virendra Singh, Prof. Sachin Patkar, and Prof. Madhava Desai.
- IIT Madras: Prof. V Kamakoti, Dr. Rupesh Nasre, Prof. Madhu Mutyam
- IIT Delhi: Prof. Pritiranjan Panda, Prof. Smruti Sarangi
- IIT Kanpur: Prof. Mainak Chaoudhuri, Dr. Biswabandan Panda
- IIT/IIIT Hyderabad: Dr. Mittal, Dr. Lavanya

Thank You

Acknowledgement:

- Prof Virendra Singh, IIT Bombay
- Prof Matthew Jacob, IISc Bangalore
- Prof R Govind Rajan, IISc Bangalore
- Prof C Chandrasekhar, IIT Madras
- Prof Kewal Saluja, Uni of Wisconsin
- Prof Adit Singh, Auburn University
- Prof Ravi Iyer, IIT Tirupati