COSC175 (Systems I): Computer Organization & Design

Professor Lillian Pentecost
Fall 2024

Warm-Up October 17

- Where we were
  - Fall Break!!
- Where we are going
  - Bringing our history tour to the 21st century!!
  - Mid-Term feedback
- Logistics, Reminders
  - TA help 7-9PM on Sundays, Tuesdays, Thursdays in C107
    - **SPECIAL midterm TA session with FOOD and EXTRA PROBLEMS, Sunday, Oct. 20 6-9PM**
  - LP Office hours M 9-10:30AM, Th 2:30-4PM
  - NO Weekly Exercises this week; see extra practice problems (will be posted tonight)
  - Lab 4 Report due October 17 10PM
  - Weekly Exercise solutions will *not* be posted, but you are encouraged to ask (and respond to!) about specific questions or confusions using the Moodle Questions forum
  - ***Midterm Exam, Oct. 22, 2 hour exam, 6-9:30PM, Pruyne Lecture Hall (show up no later than 7:30!!)***

Midterm Logistics (any questions?)

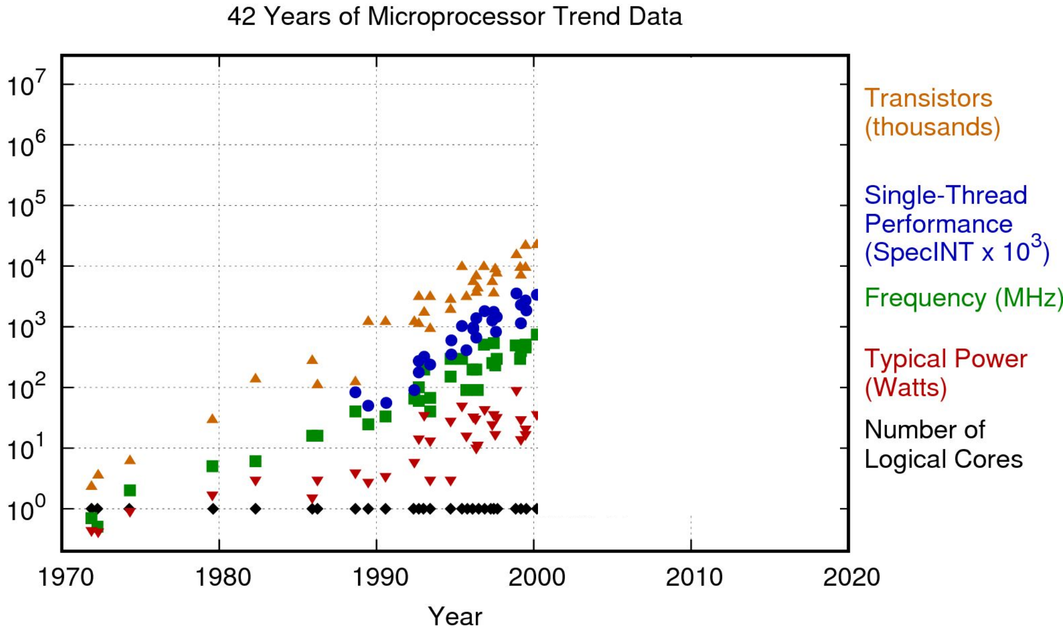
- Basics
    - You can type or handwrite; you will hand in a hard-copy, and optionally upload to Moodle
    - Tuesday, October 22, 2 hour exam + 5 minutes to submit via Moodle
    - ***You can use any of your own notes, and any materials on our Moodle page, no external resources***
    - I will grade ***anonymously*** (you will be assigned a pseudonym and given instructions when you arrive; you will make sure your pseudonym is on EVERY PAGE you submit)
- What types of questions? 1 short answer section + 3 “analysis+design” sections
    - Short-answer (2-3 sentences), from assigned readings and all in-class material (through today), these will be similar to textbook exercises and especially the “Interview Questions” at the end of each section
    - “Analysis and Design” sections will be multi-part problems, more similar to check-ins and lab reflection questions
        - Be able to translate a plain-text description of a system’s behavior into a circuit diagram
        - Be able to analyze a provided SystemVerilog implementation of a combinational or sequential logic design (you will not be asked to write any SystemVerilog from scratch)
        - Demonstrate comfort with binary representation, boolean expressions, and boolean algebra
        - Demonstrate understanding of important digital building blocks (e.g., ALU, Memory Arrays), and be able to analyze variations and extensions of these building blocks.
- How to study?
    - Review readings, weekly exercises, notes, and **lab reflection questions**
    - Make sure you understand terms and check-in questions from throughout the semester

Our Brief Computer History Tour: Some Caveats

- This will be a USA-centric view, at least to start
- We will move quickly, and will focus on how the resources and the abilities of computers have evolved over the last century or so
- There are countless social, economic, and political questions to keep in mind
  - Why are these machines being built?
  - Where are these machines being developed, and who has access to them?
  - Who is doing the work of computing at each of these points in time?
- https://www.computerhistory.org/ has some great articles and is the main source of the dates and details I'll include today
- All images are gathered either via CHM or CC
- IBM archives are also an interesting resource

Banking in on Moore's Law: 1980s-1990s

- Machines get faster, memory gets denser, capabilities increase!
- Let the good times roll!
- Individual transistor dimensions scale from 40 microns to (most recently) ***<u>3-7 nanometers</u>***



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

Banking in on Moore's Law: 1980s-1990s

- 16 MHz+ CPU
- 4MB+ RAM
- A Hard Drive! Maybe 160MB
- Packaging a display, pre-loading some software

A new key issue emerges: how to quickly and easily share data, software across machines?

[ARPAnet]

Ordering your AOL installation CD in the mail to help you get connected online – early messaging and email

[Listen to Steve Case of AOL on `How I Built This`, NPR]

Emergence of the Memory Wall

Hitting the Memory Wall: Implications of the Obvious

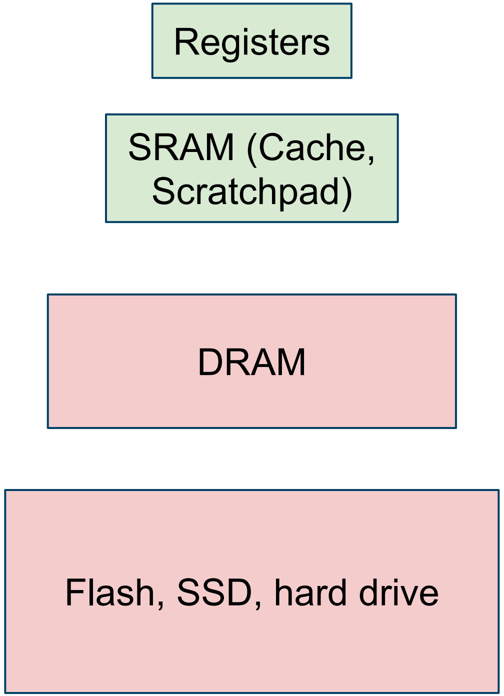
Wm. A. Wulf
Sally A. McKee
Department of Computer Science
University of Virginia
{wulf | mckee}@virginia.edu

December 1994

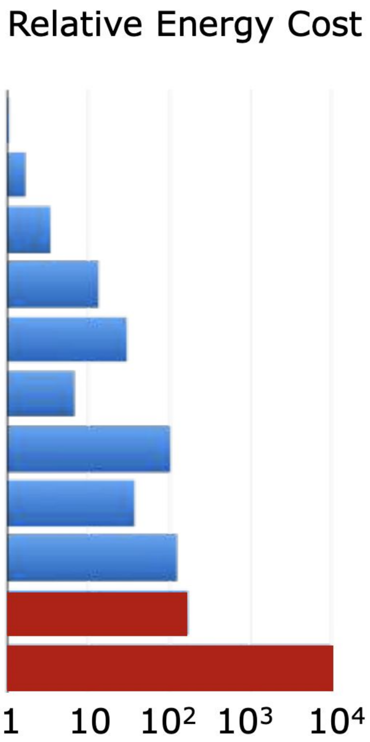
This brief note points out something obvious — something the authors "knew" without really understanding. With apologies to those who did understand, we offer it to those others who, like us, missed the point.

As noted above, the right solution to the problem of the memory wall is probably something that we haven't thought of — but we would like to see the discussion engaged. It would appear that we do not have a great deal of time.

Accessing memory remains strictly more costly; what does this mean for our applications?



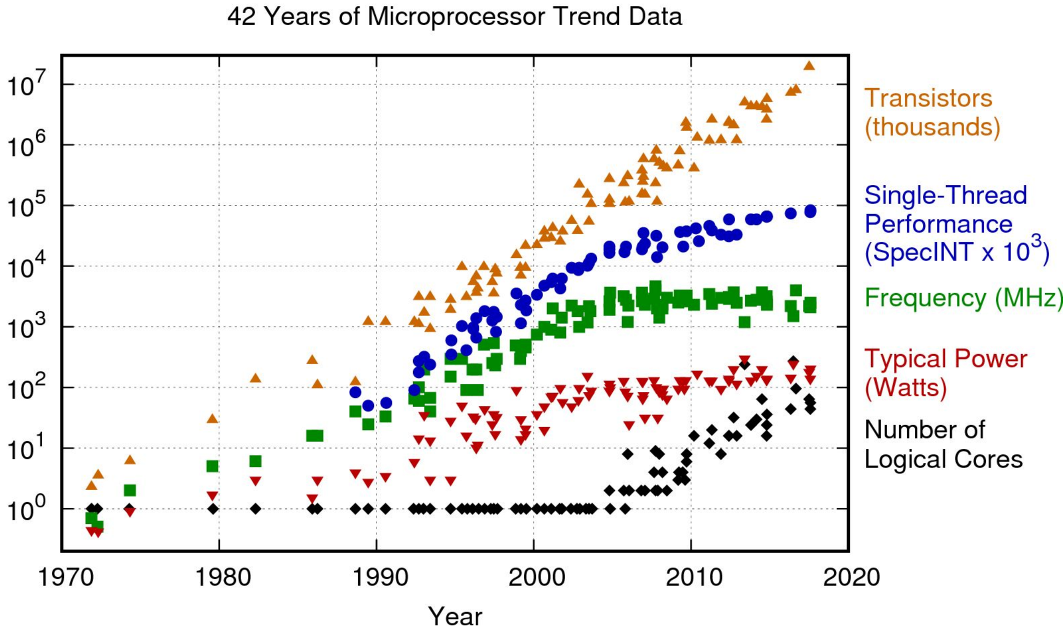
Operation:	Energy (pJ)
8b Add	0.03
16b Add	0.05
32b Add	0.1
16b FP Add	0.4
32b FP Add	0.9
8b Multiply	0.2
32b Multiply	3.1
16b FP Multiply	1.1
32b FP Multiply	3.7
32b SRAM Read (8KB)	5
32b DRAM Read	640



[**Horowitz**, *ISSCC* 2014]

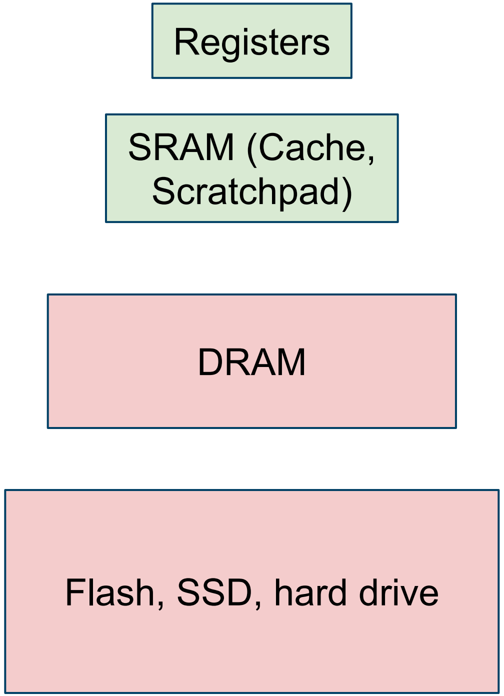
Multi-Core; Delaying the Inevitable: 2000s

- Increase number of logical cores – do more in parallel to maintain same amount of productivity
- Power density and ***dark silicon***
- Complexity and management of devices increases, ***programmability and usability are paramount***

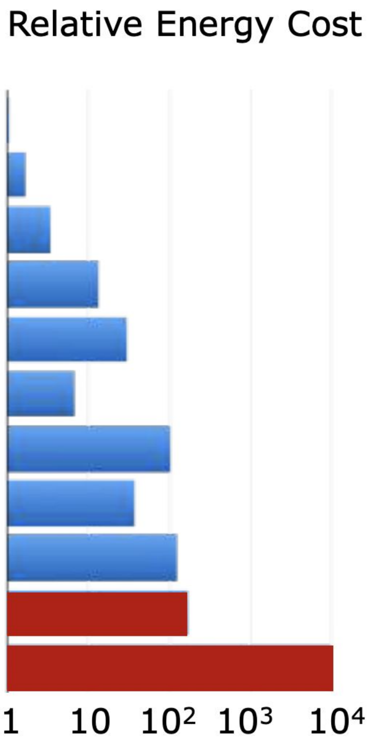


Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

Accessing memory remains strictly more costly; what does this mean for our applications?



Operation:	Energy (pJ)
8b Add	0.03
16b Add	0.05
32b Add	0.1
16b FP Add	0.4
32b FP Add	0.9
8b Multiply	0.2
32b Multiply	3.1
16b FP Multiply	1.1
32b FP Multiply	3.7
32b SRAM Read (8KB)	5
32b DRAM Read	640



[**Horowitz**, *ISSCC* 2014]

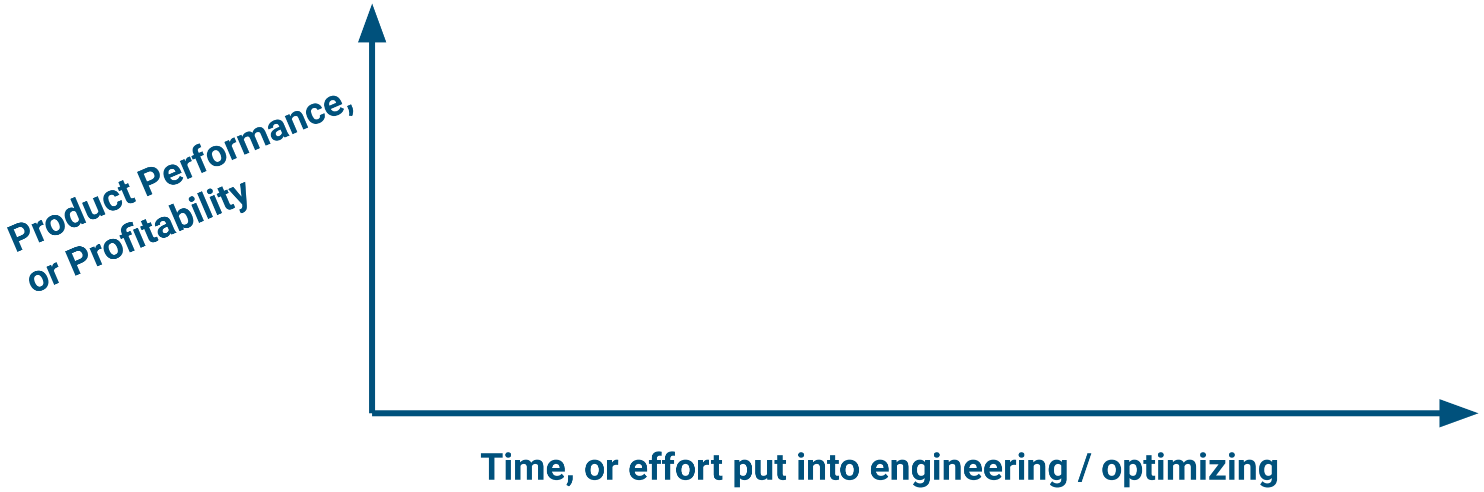
Demo time! If you aren't at the demo, you should be filling out Midterm Feedback Form

End of Moore’s Law; Modern Processors

- S-Curves of Technology → What does it mean for Moore’s Law to end?
  - *Physical* limitations vs. *financial* limitations

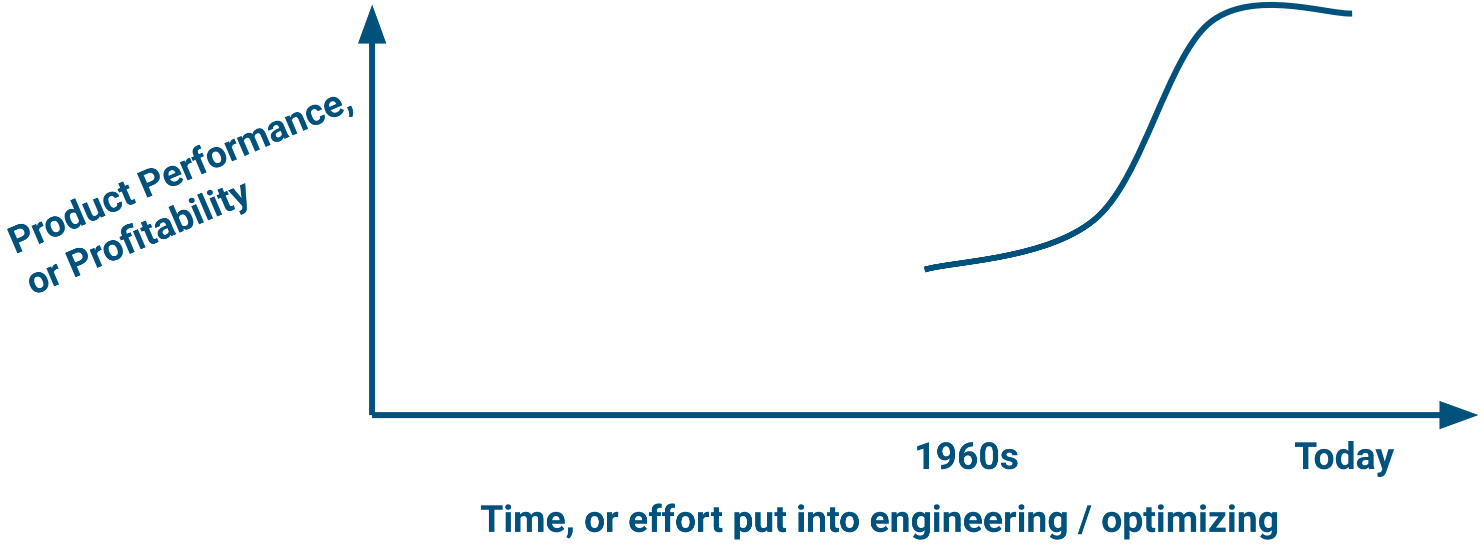
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



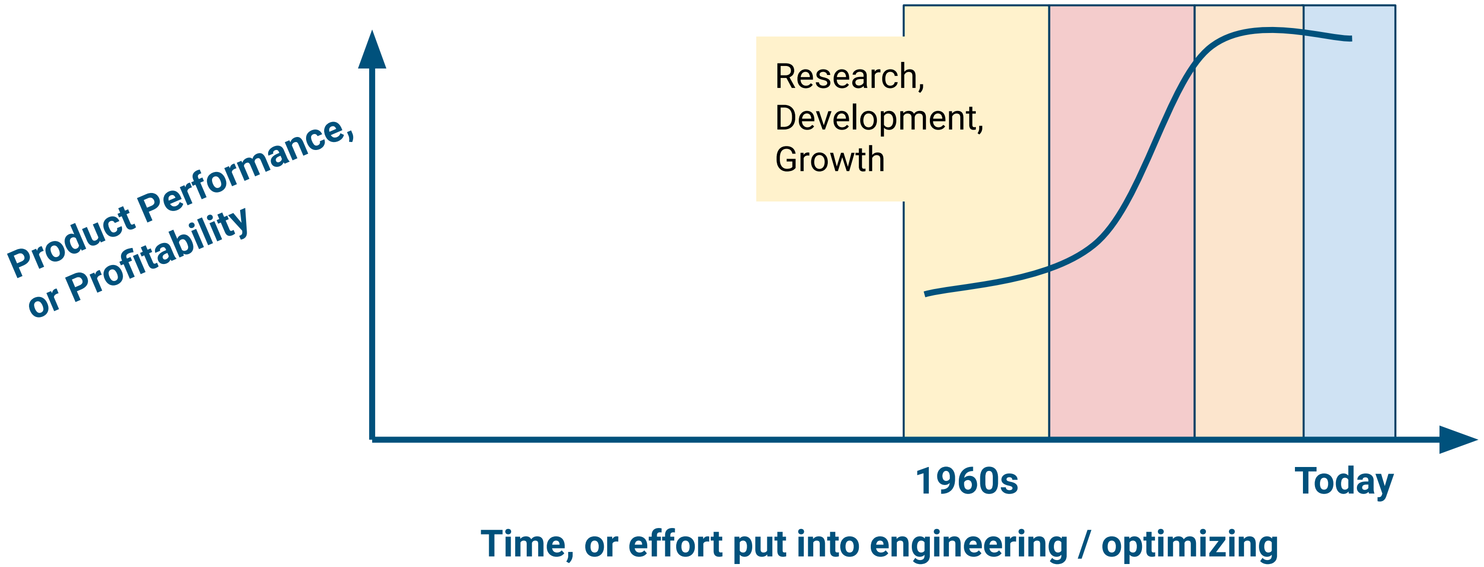
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



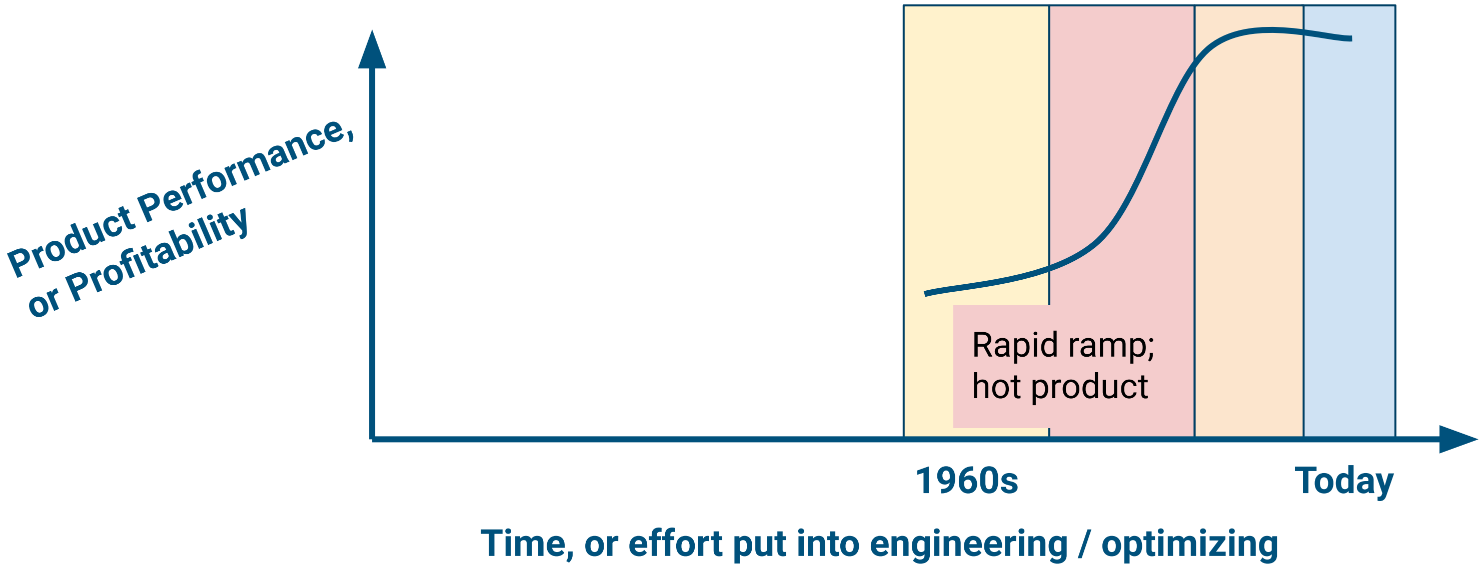
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



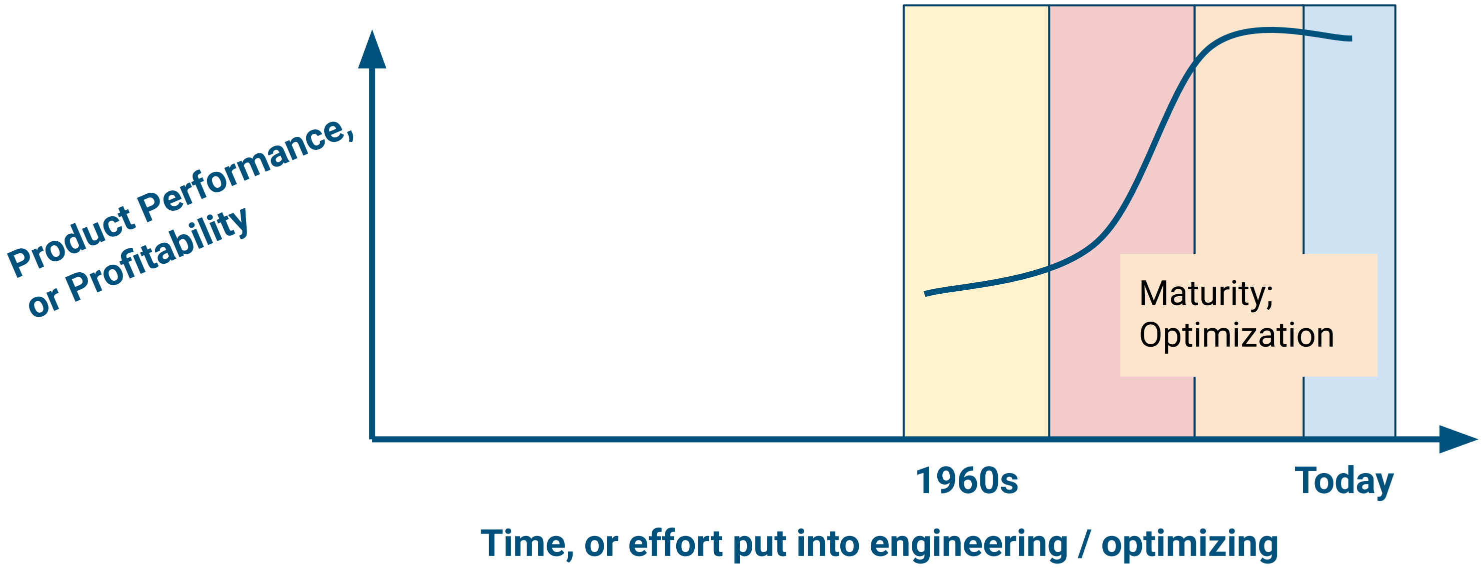
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore’s Law to end?
    - *Physical* limitations vs. *financial* limitations



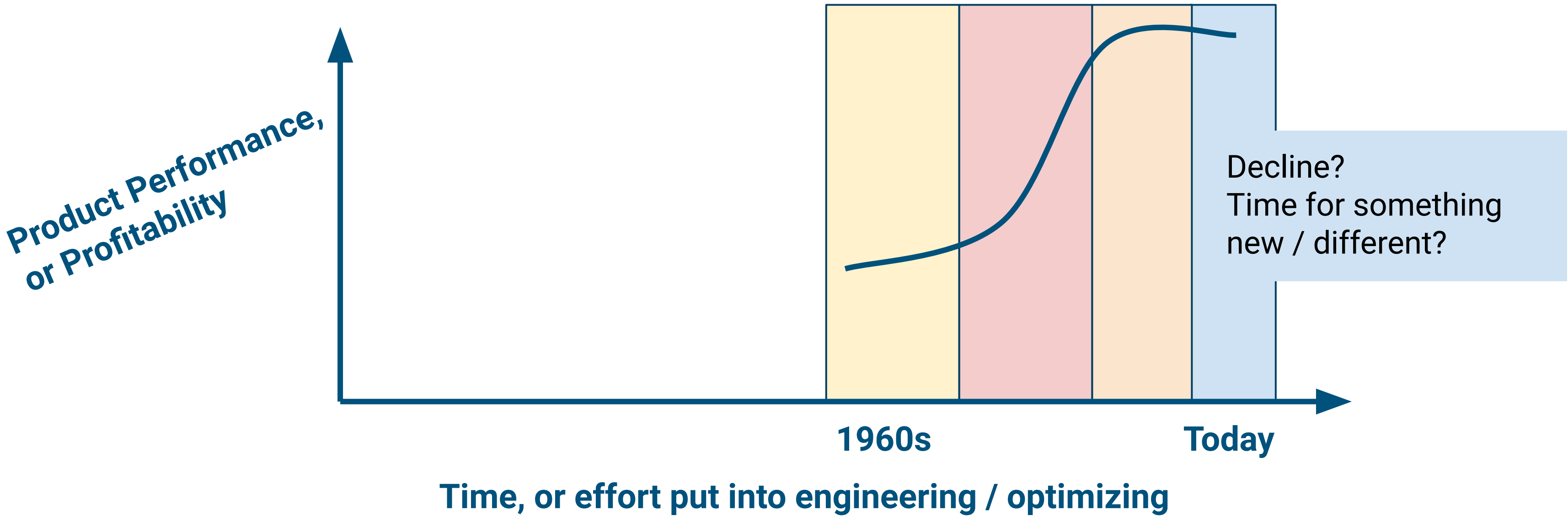
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



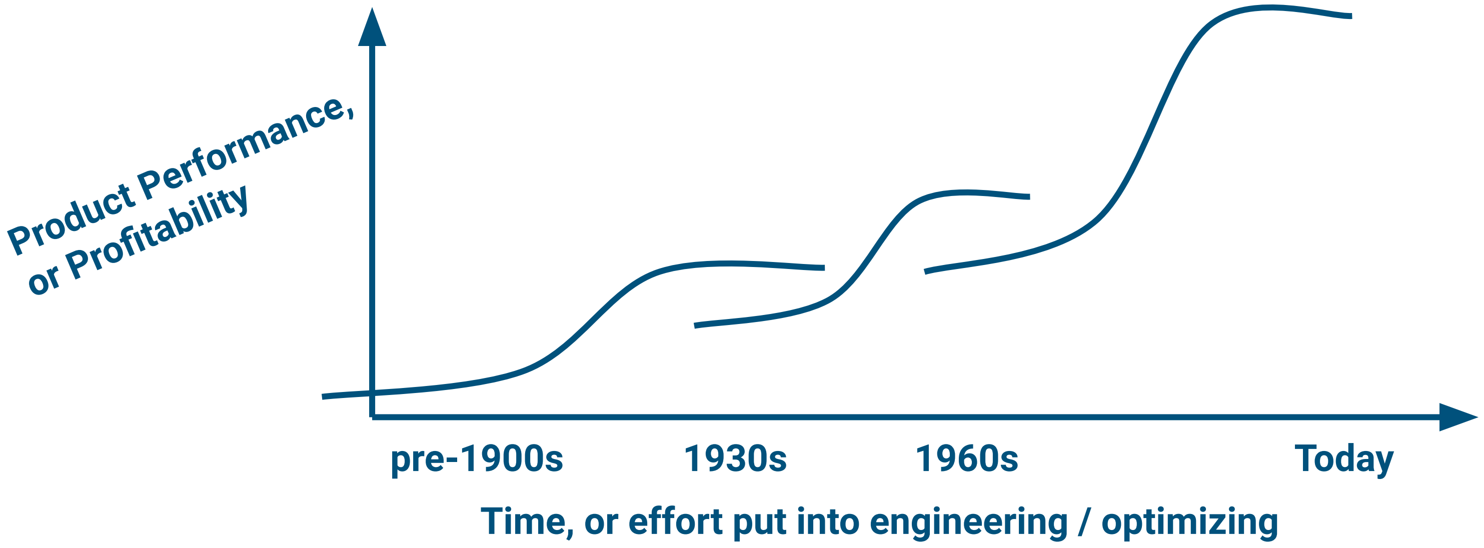
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



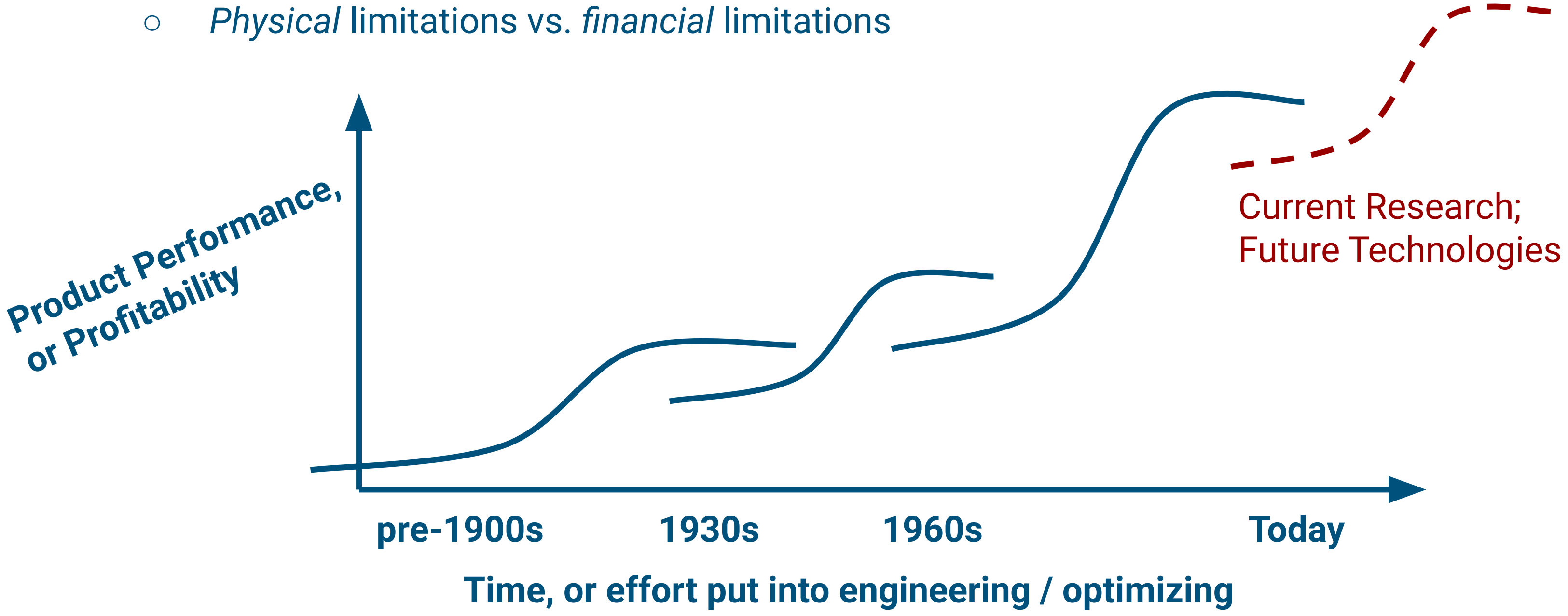
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



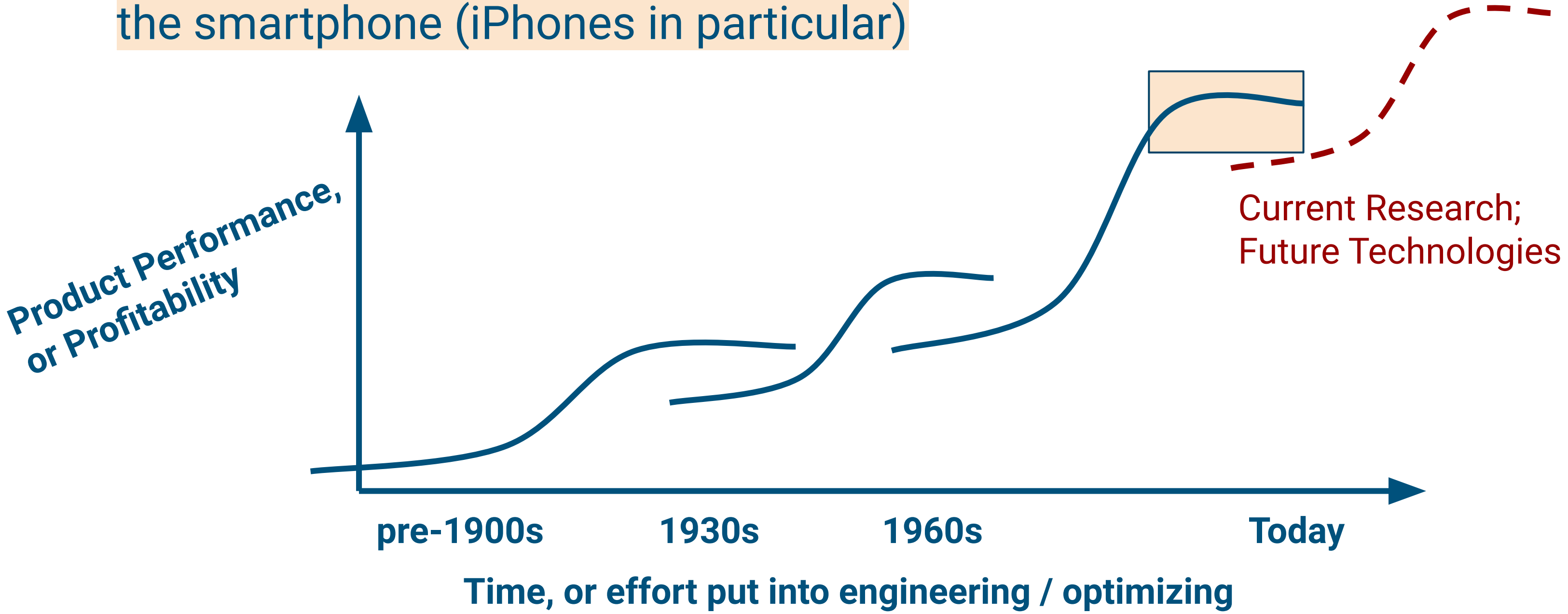
Thinking about trends (not a real graph)

- S-Curves of Technology → What does it mean for Moore's Law to end?
  - *Physical* limitations vs. *financial* limitations



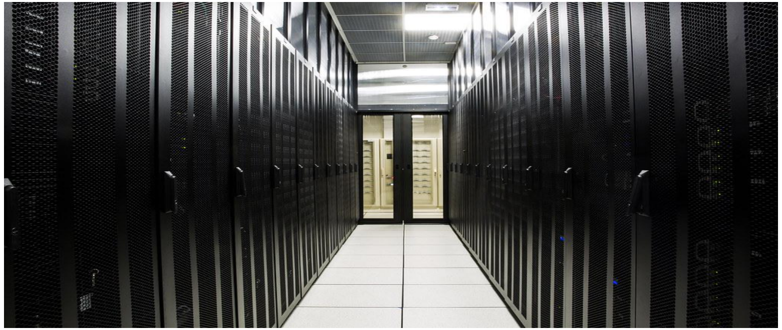
Thinking about trends (not a real graph)

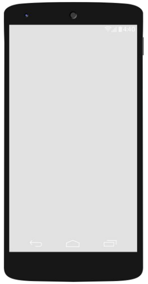
- Let's take as an example, 15 years in the development and optimization of the smartphone (iPhones in particular)



End of Moore's Law; Modern Processors

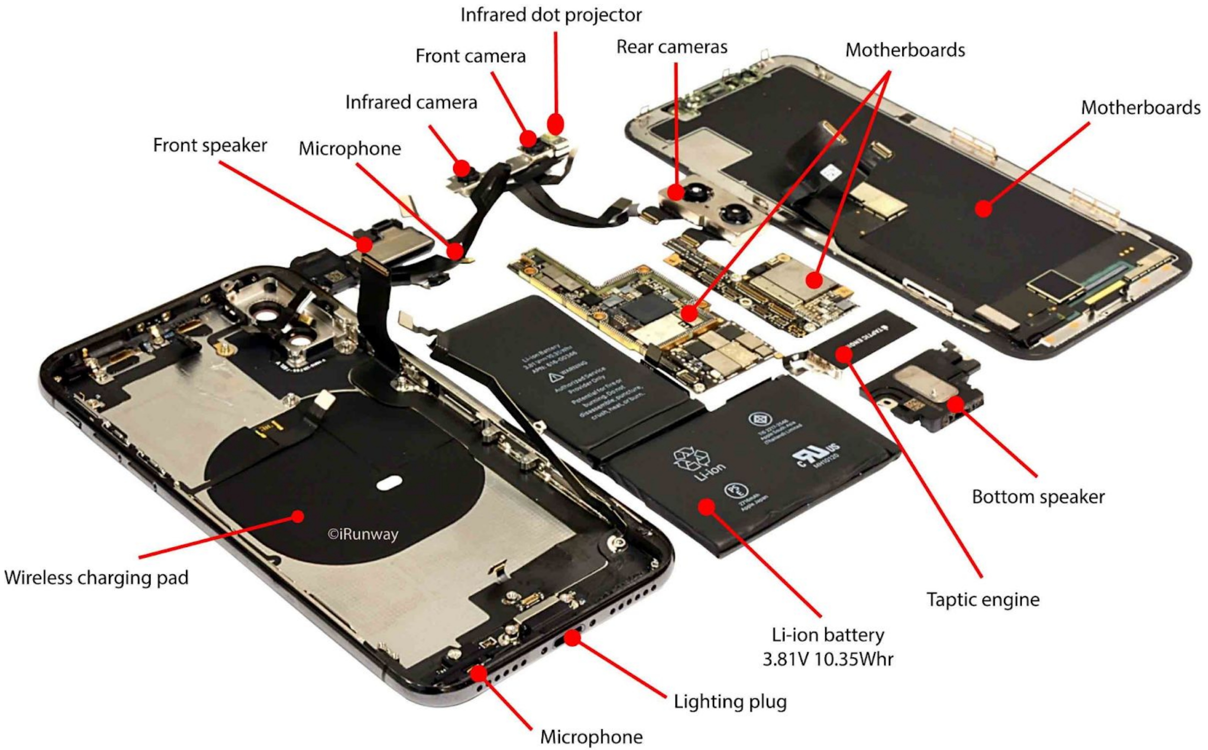
Capabilities vary with application and scale





Anatomy of a Smartphone

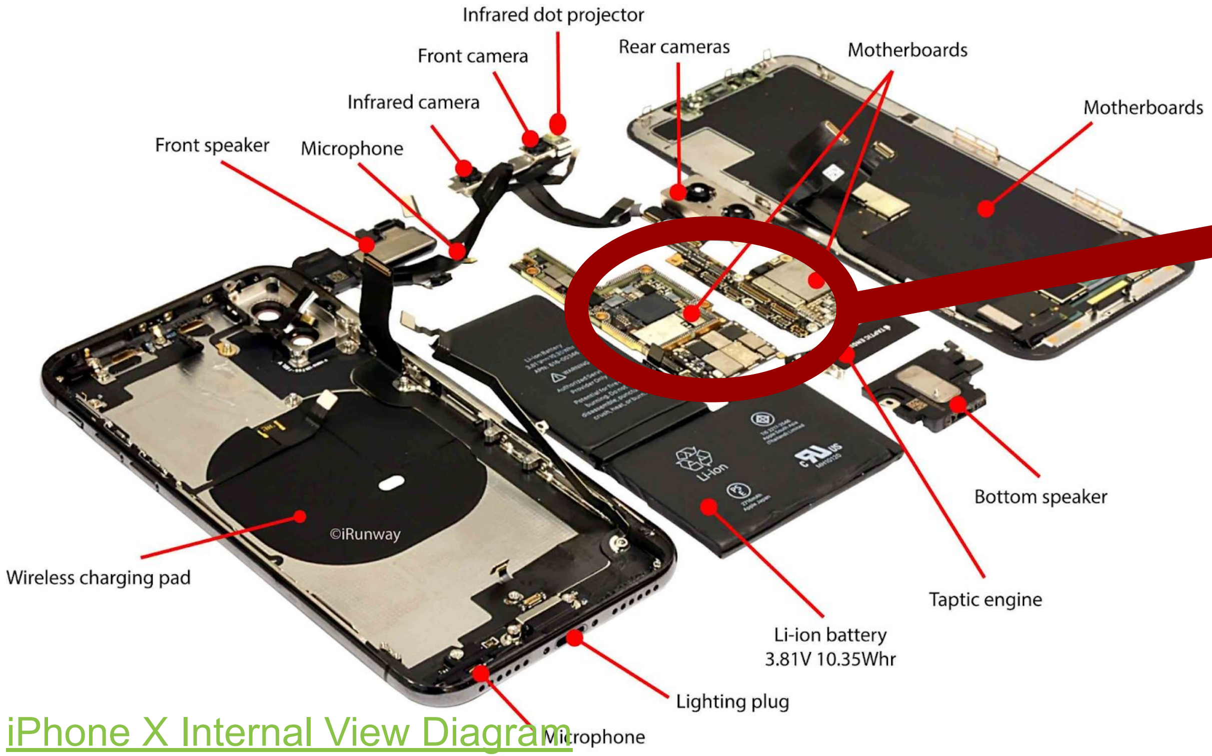
- What are the limitations? Where is on-chip real estate spent?



[iPhone X Internal View Diagram]

Anatomy of a Smartphone

- What are the limitations? Where is on-chip real estate spent?

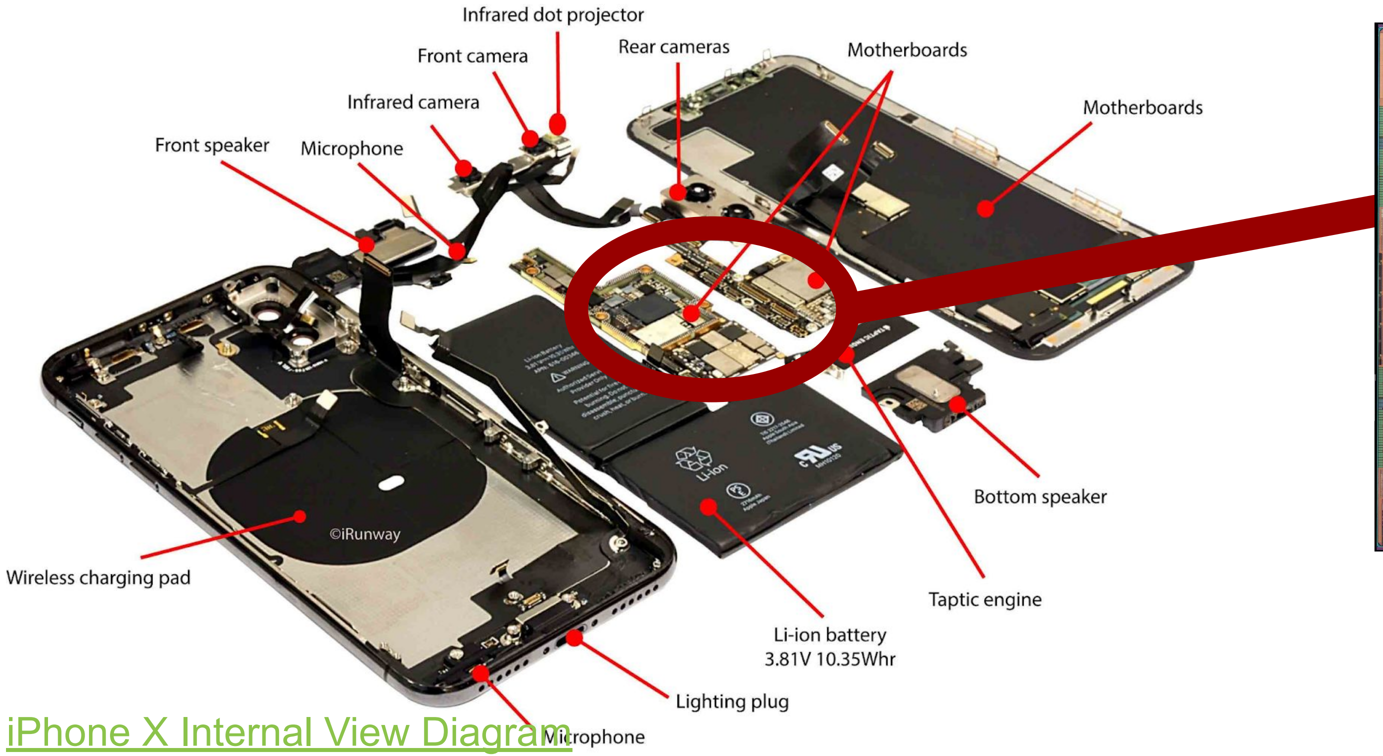


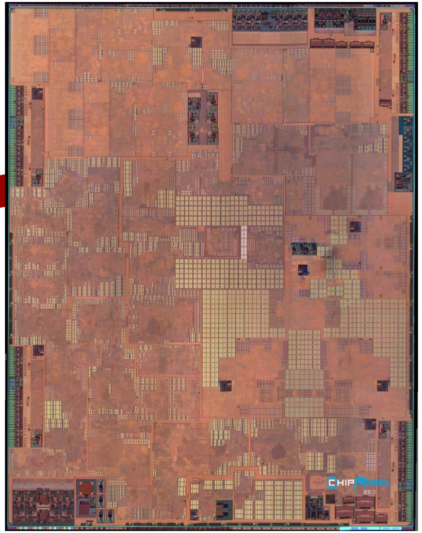
A11 Chip Views

iPhone X Internal View Diagram

Anatomy of a Smartphone

- What are the limitations? Where is on-chip real estate spent?



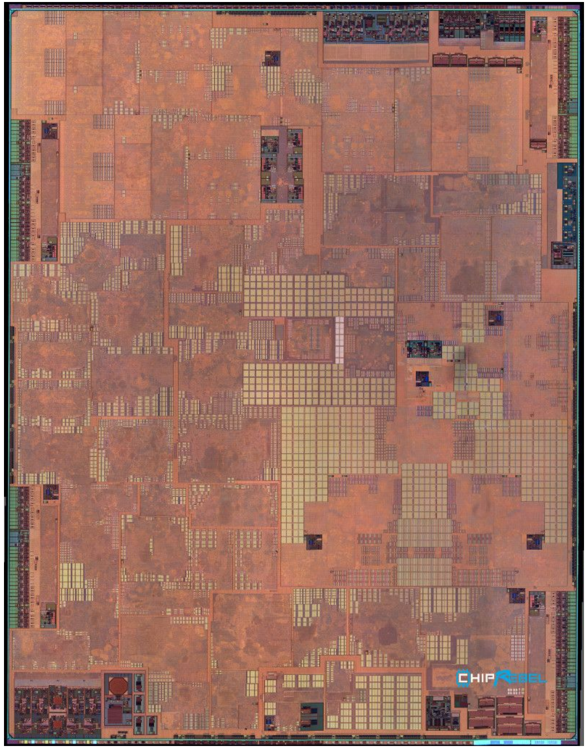


A11 Chip Views

iPhone X Internal View Diagram

Anatomy of a Smartphone

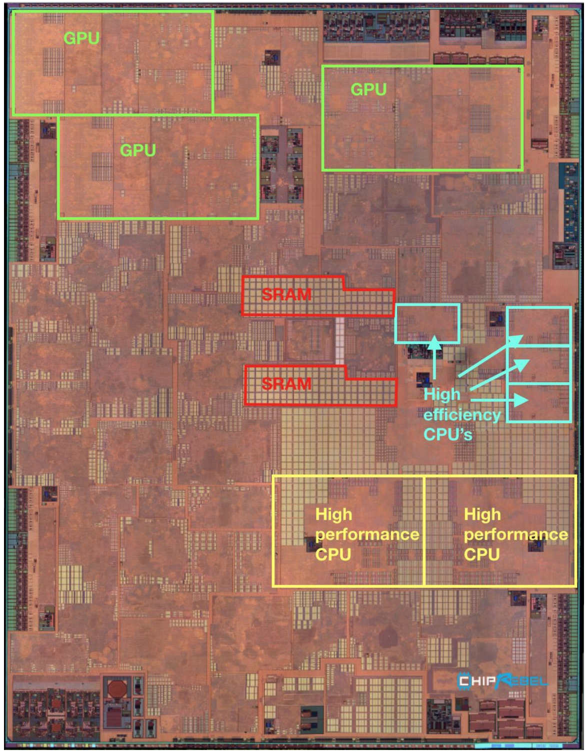
- What are the limitations? Where is on-chip real estate spent?



[A11 Chip Views]

Anatomy of a Smartphone

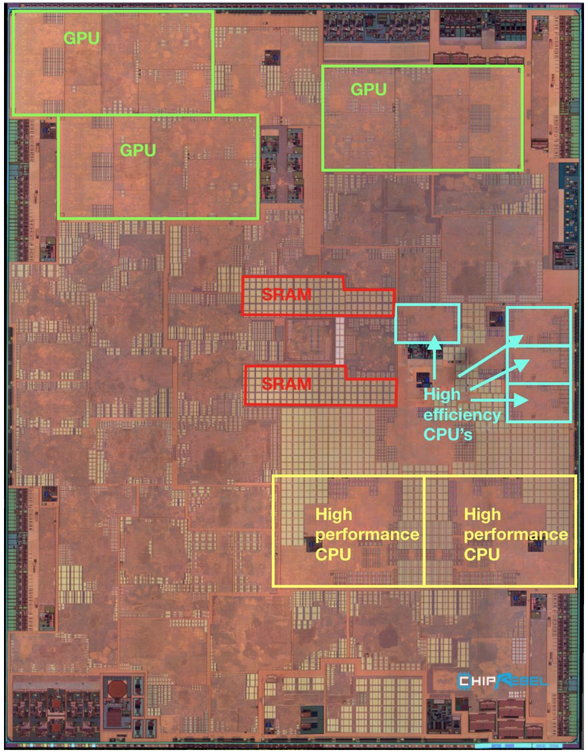
- What are the limitations? Where is on-chip real estate spent?



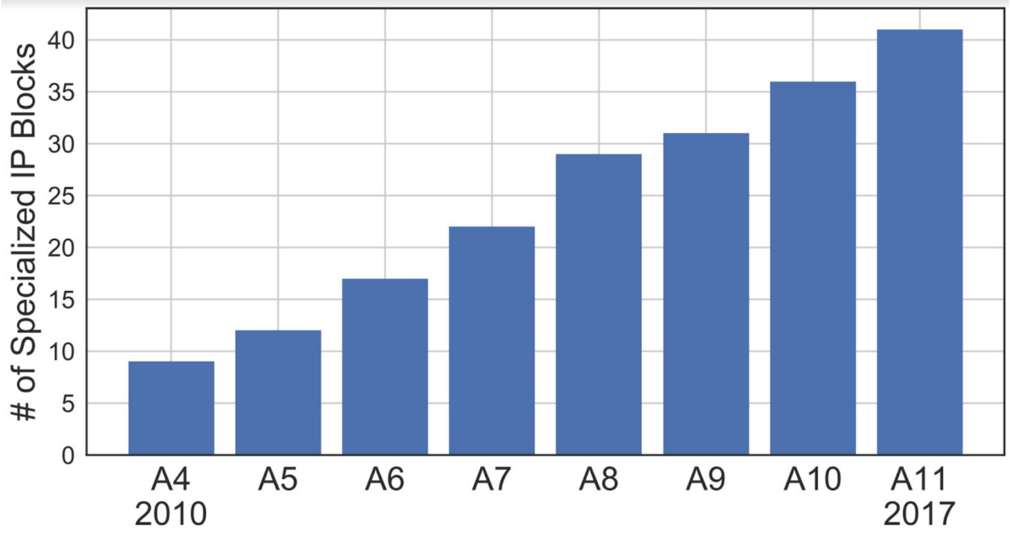
[A11 Chip Views]

Anatomy of a Smartphone

- What are the limitations? Where is on-chip real estate spent?

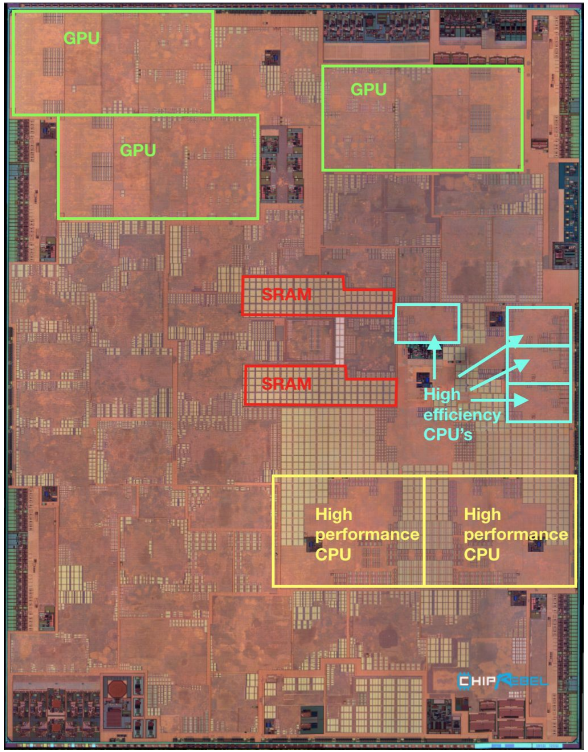


A11 Chip Views



Die Photo Accelerator Analysis

Accessing memory remains strictly more costly; what does this mean for our applications?

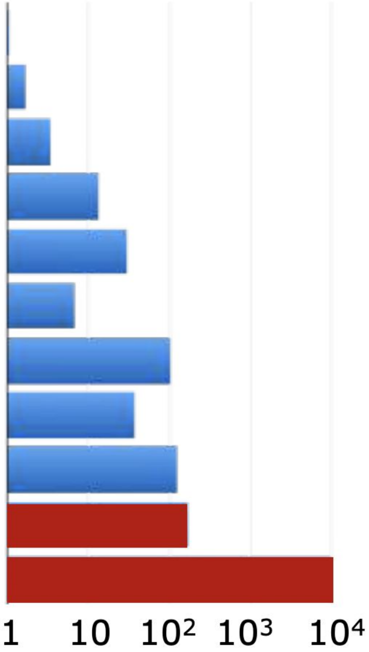


A11 Chip Views

Operation:	Energy (pJ)
8b Add	0.03
16b Add	0.05
32b Add	0.1
16b FP Add	0.4
32b FP Add	0.9
8b Multiply	0.2
32b Multiply	3.1
16b FP Multiply	1.1
32b FP Multiply	3.7
32b SRAM Read (8KB)	5
32b DRAM Read	640

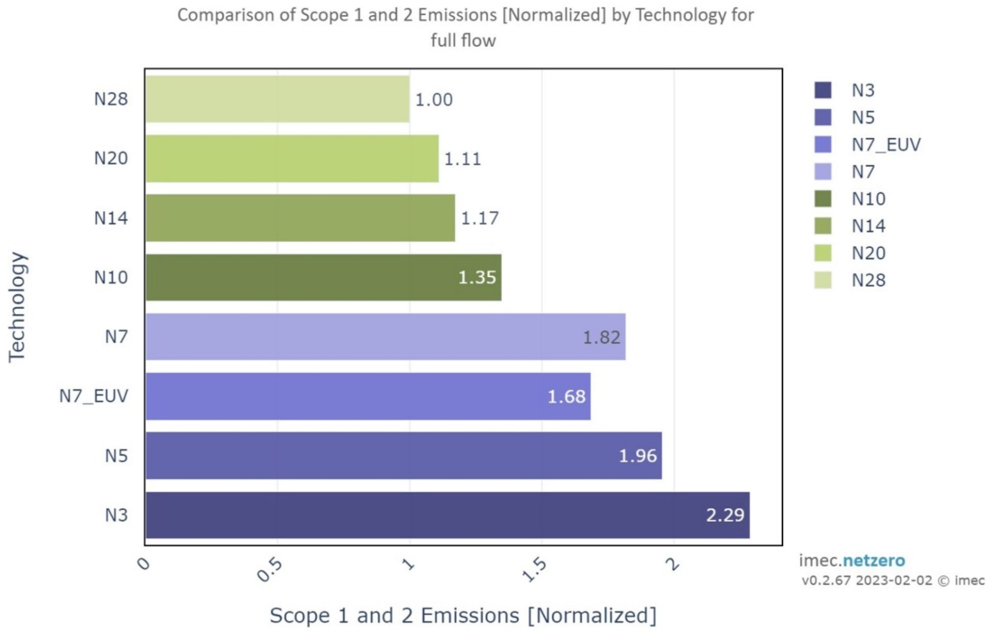
[**Horowitz**, *ISSCC* 2014]

Relative Energy Cost



Smaller and smaller transistors: what are the other costs?

- Individual transistor dimensions scale from 40 microns to (most recently) ***3-7 nanometers***
- Energy efficiency is critical, but is not the main contributor to the ***environmental footprint*** of technology



Bring it all together: what's really in a modern computer

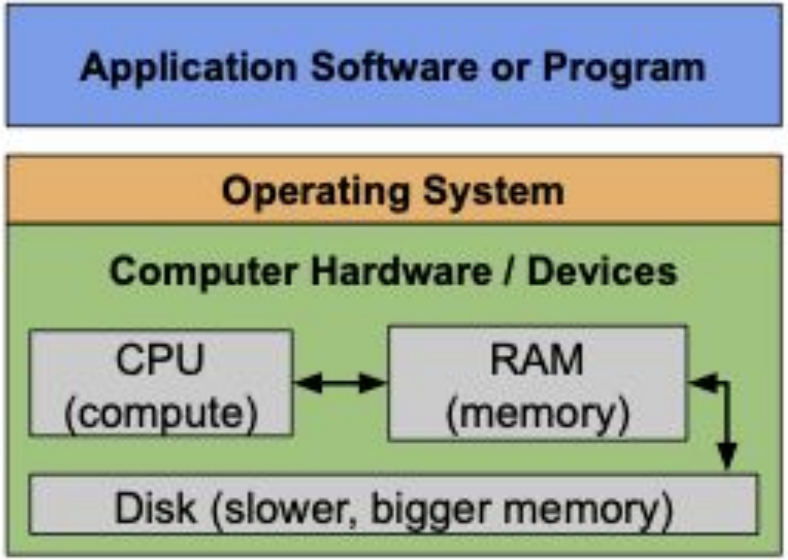
– How has each layer of the computer system evolved over the past 50 years?

<u>– Interfaces</u>

<u>– Capabilities</u>

<u>– Implementations</u>

We've fleshed out our ***hardware building blocks***, now we'll think about how to organize and use these blocks to ***execute real programs***.



Wrap-Up October 17

SCAN ME

- Coming up next!
    - Midterm exam, then switching to how to interact with HW as a programmer!!
- Logistics, Reminders
    - TA help 7-9PM on Sundays, Tuesdays, Thursdays in C107
        - **SPECIAL midterm TA session with FOOD and EXTRA PROBLEMS, Sunday, Oct. 20 6-9PM**
    - LP Office hours M 9-10:30AM, Th 2:30-4PM
    - Weekly Exercise solutions will *not* be posted, but you are encouraged to ask (and respond to!) about specific questions or confusions using the Moodle Questions forum
    - You can expect Lab & Weekly Exercise feedback to be posted by Sunday, Oct. 20
    - ***Midterm Exam, Oct. 22, 2 hour exam, 6-9:30PM, Pruyne Lecture Hall (show up by 7:30!!)***
- FEEDBACK
    - https://forms.gle/5Aafcm3iJthX78jx6

Moore's Law Check-In

(1)Read just the first section of original Moore's Law paper aloud at your table (until "Past and future", or as far as you get, see PDF on Moodle),
(2) talk to your neighbors about what surprises you, what confuses you,
(3) write down a question