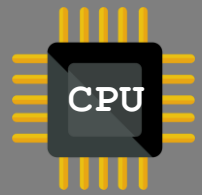


Topic 2

Quantitative Design and Analysis #2

Xinfei Guo
xinfei.guo@sjtu.edu.cn

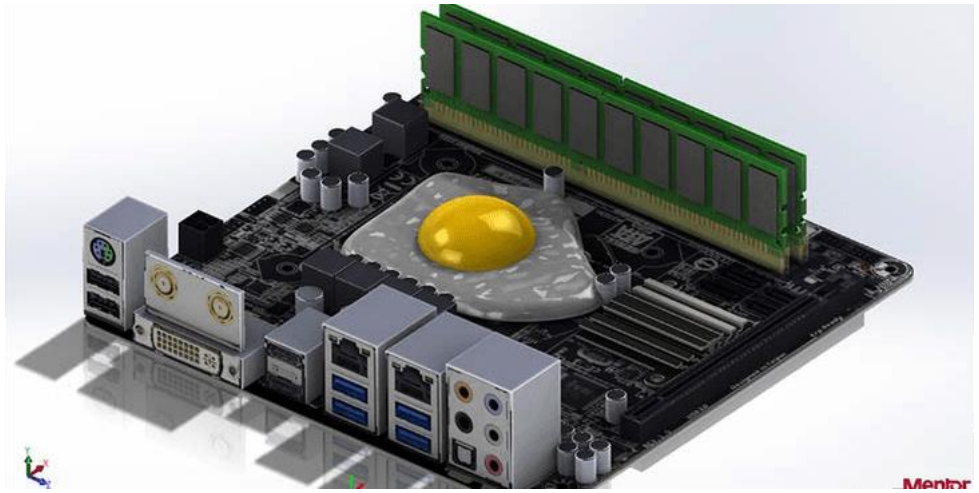
May 16th, 2022

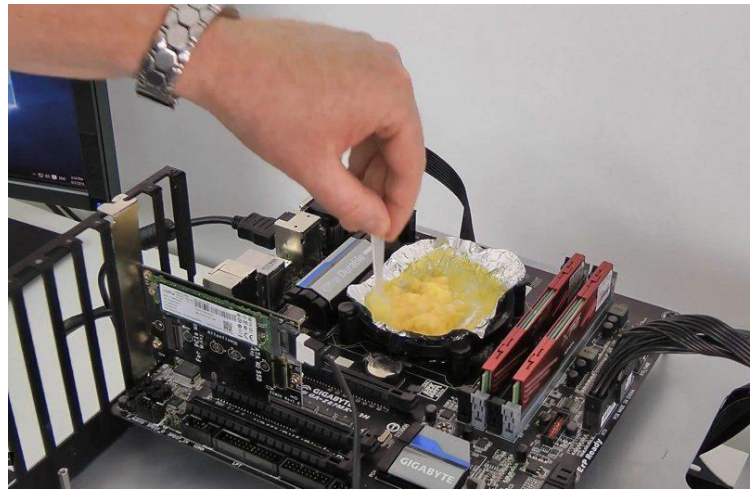
T2 learning goals

- **Quantitative** Design and Analysis
  - Section #1
    - ISAs
    - Performance
  - Section #2
    - Power and Energy
    - Reliability
    - Cost

Why power is important?

- Imagine how hot is a chip?
  - The CPU junction temperature exceeds 90°C within 6 seconds…
  - Cooling is expensive!





Can You Really Fry An Egg On A CPU?

source:

- https://semiengineering.com/can-you-really-fry-an-egg-on-a-cpu/
- https://www.eteknix.com/can-make-scrambled-eggs-cpu/

Trends in cooling

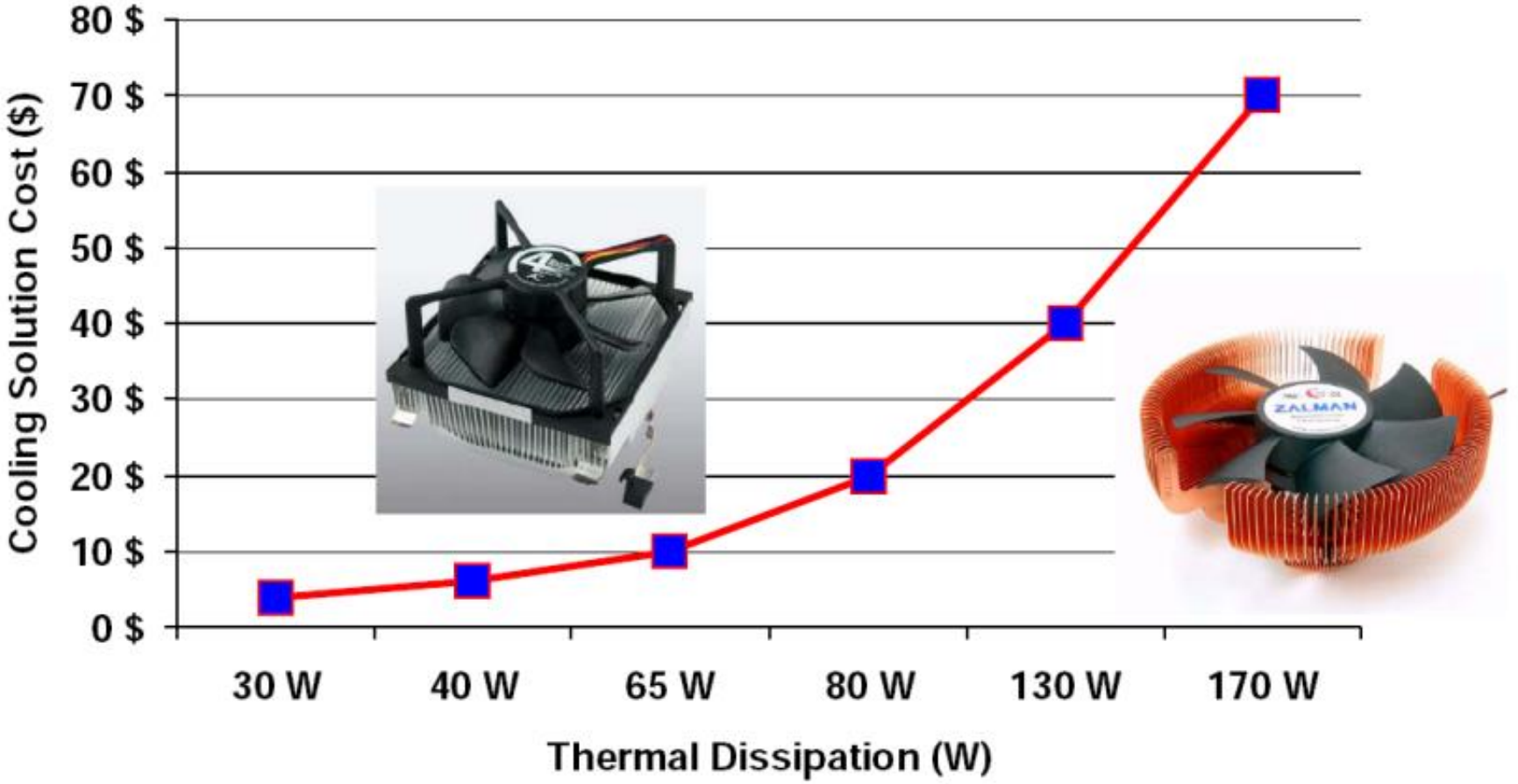
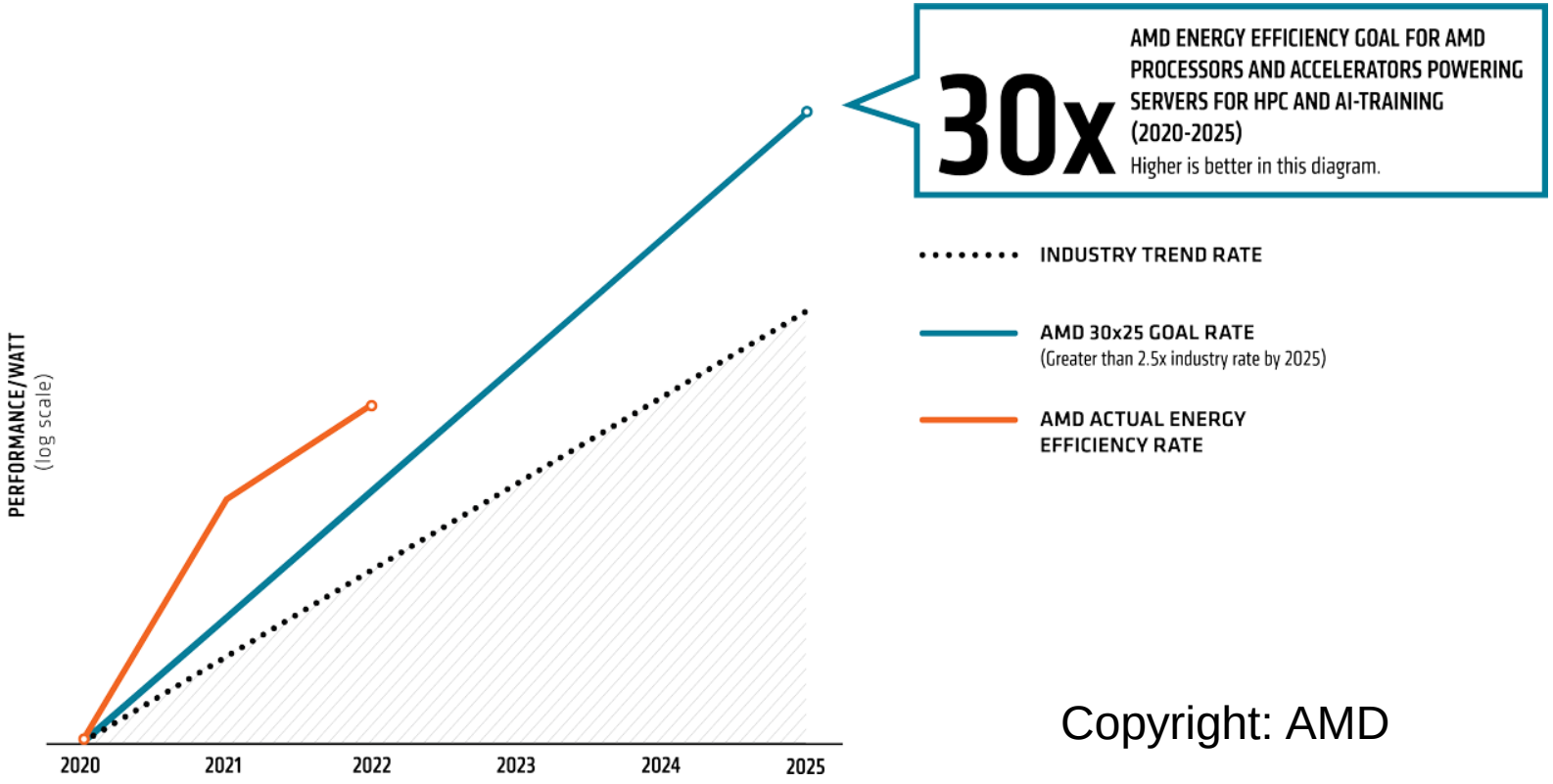


Figure: https://www.cse.wustl.edu/~lu/cse467s/slides/power.pdf

Why power is important?

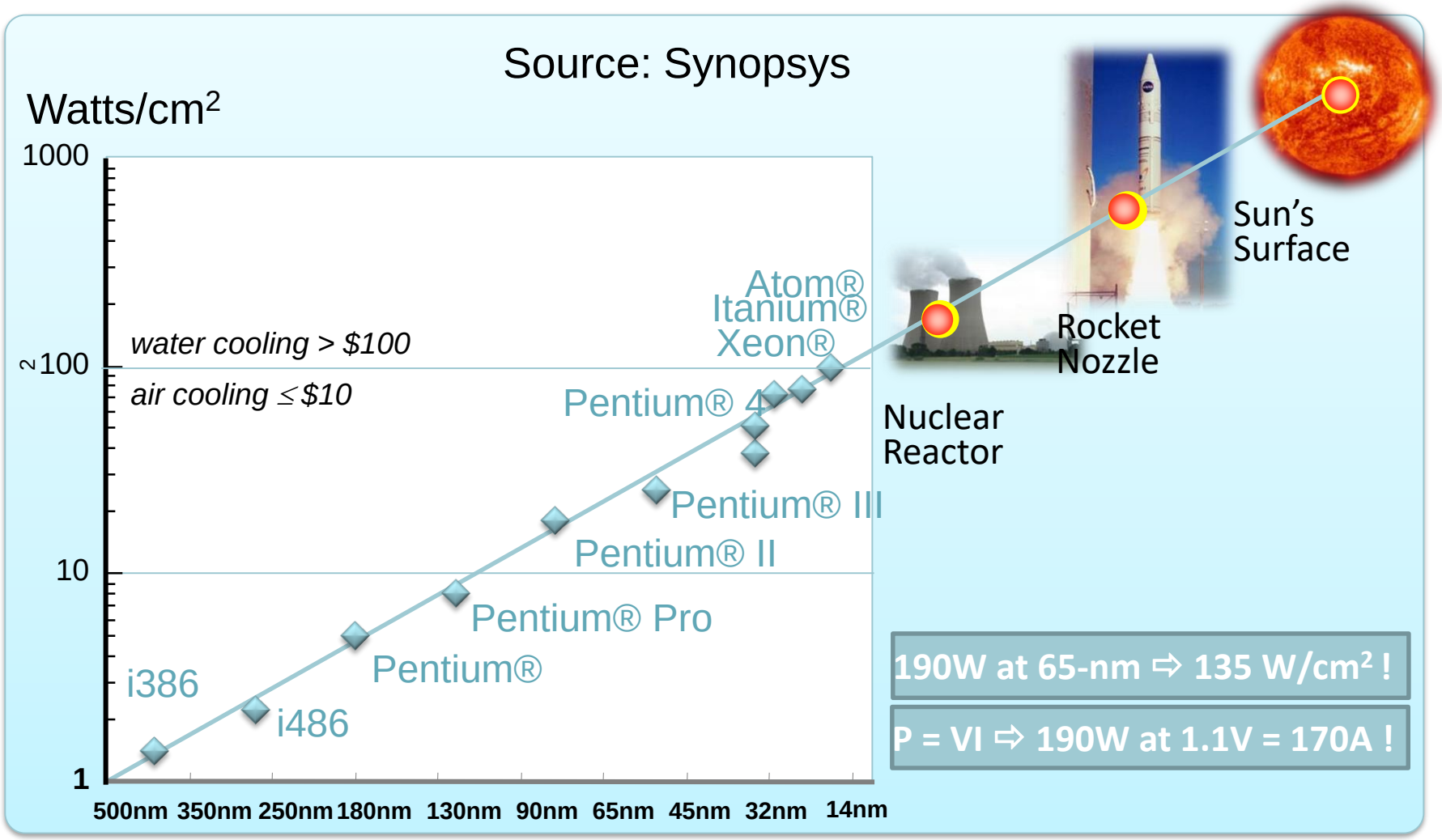
- The **30**x goal would save billions of kilowatt hours of electricity in 20**25**, reducing the power required for these systems to complete a single calculation by **97%** over five years.



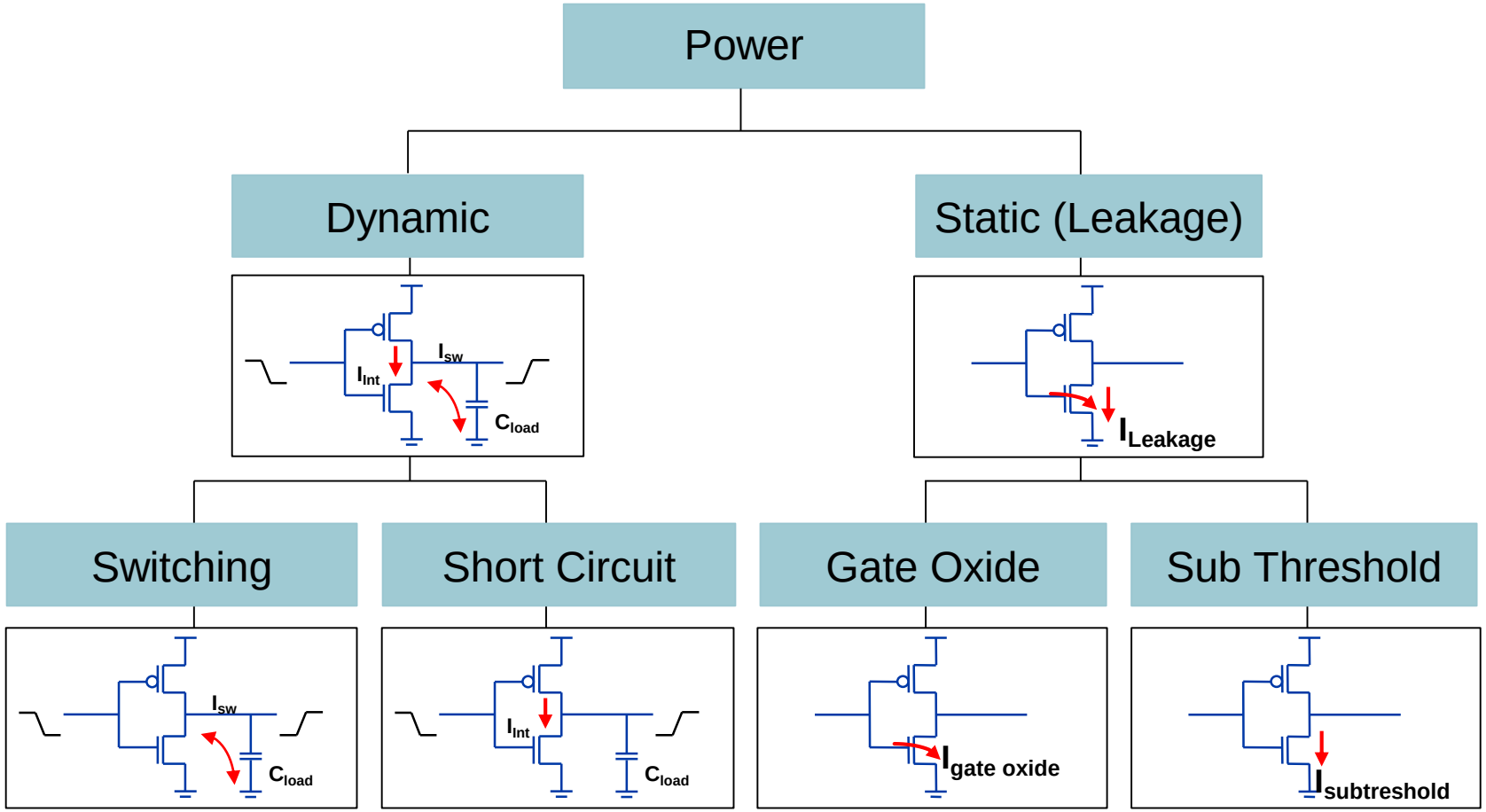
source:

- https://www.amd.com/en/corporate-responsibility/data-center-sustainability
- https://www.amd.com/en/press-releases/2021-09-29-amd-announces-ambitious-goal-to-increase-energy-efficiency-processors

Power Consumption Trends



Sources of Power



Source: Synopsys

Power Equation

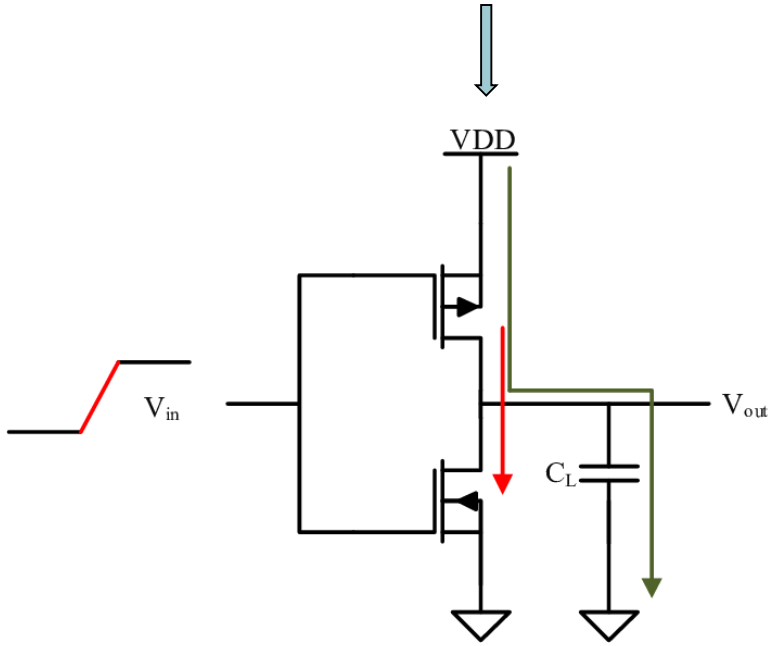
1/2 typically, but not always

$$P = \alpha.f.C_L.V_{DD}^2 + f.I_{peak}.V_{DD} + V_{DD}.I_{static}$$

Dynamic power

Short-circuit power

Static power



Borrowed from Xinfei Guo@JI, ECE4810J, 21FA

Dynamic Power

Capacitance:
Function of fan-out, wire length, transistor sizes

Supply Voltage:
Has been dropping with successive generations

$$P_{dyn} = C_L V_{DD}^2 \alpha f$$

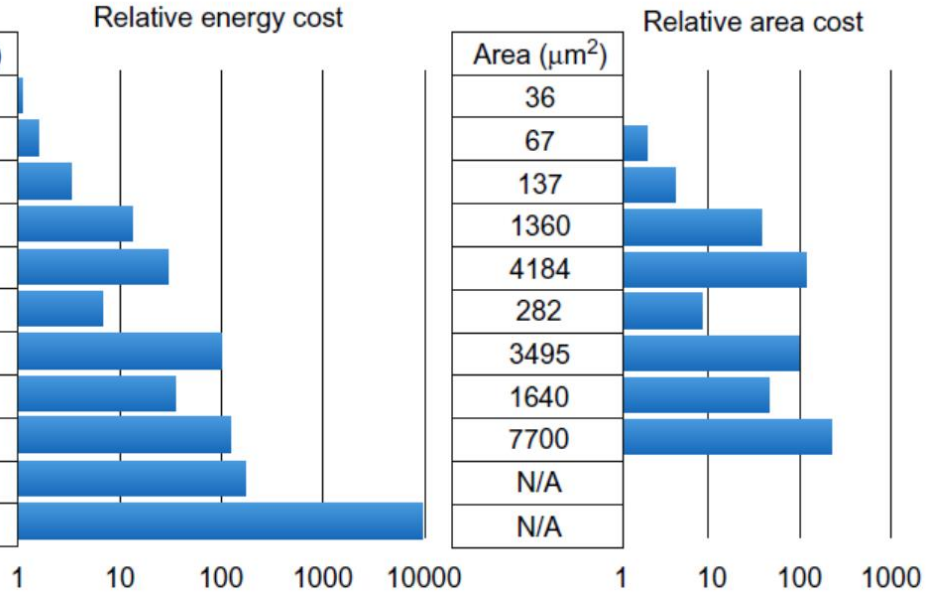
Activity factor:
How often, on average, do wires switch?

Clock frequency:
Increasing…

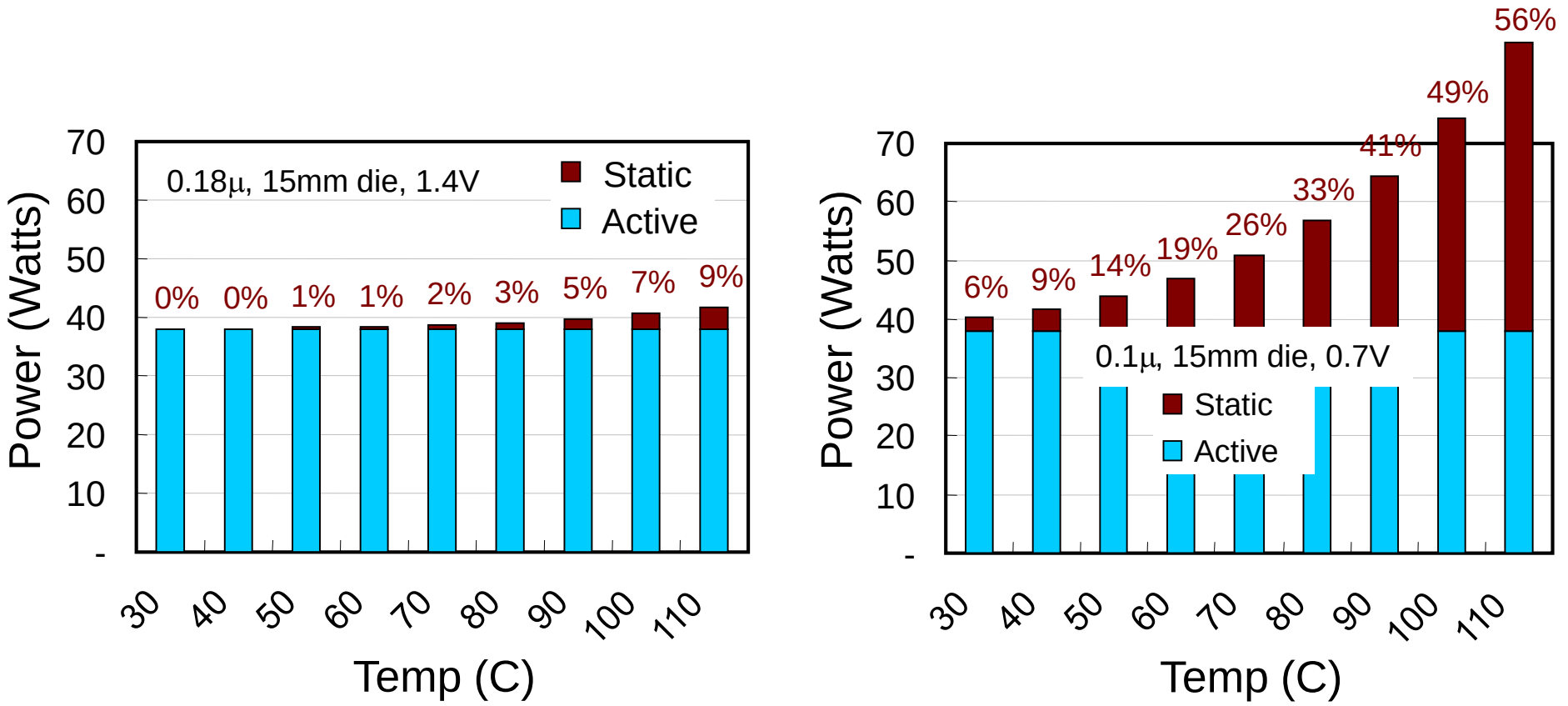
Static Power

- Static power consumption
  - 25-50% of total power
  - $Current_{static}$ x Voltage
  - Scales with number of transistors
  - To reduce: power gating

Operation:	Energy (pJ)	Area (μm^2)
8b Add	0.03	36
16b Add	0.05	67
32b Add	0.1	137
16b FB Add	0.4	1360
32b FB Add	0.9	4184
8b Mult	0.2	282
32b Mult	3.1	3495
16b FB Mult	1.1	1640
32b FB Mult	3.7	7700
32b SRAM Read (8KB)	5	N/A
32b DRAM Read	640	N/A



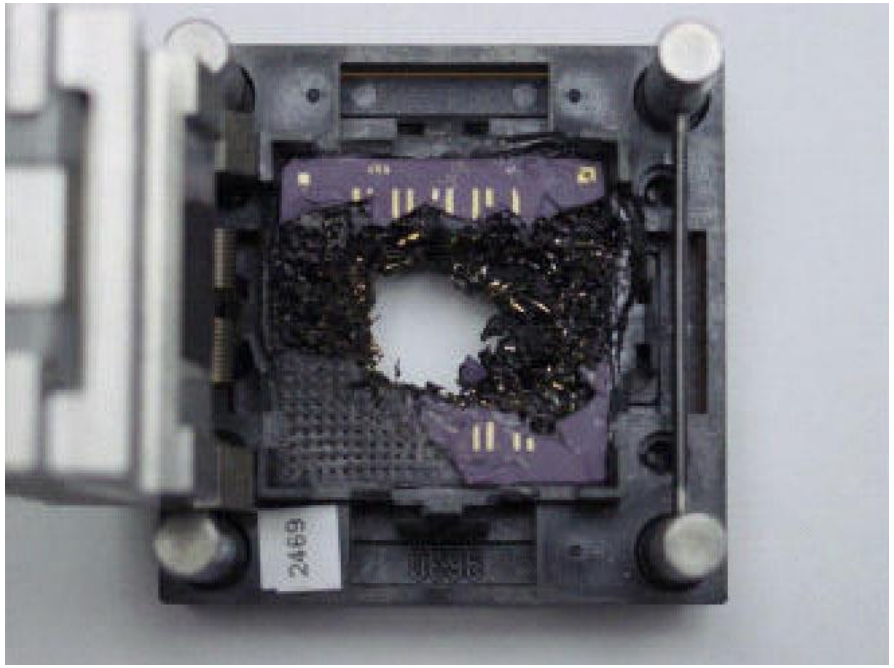
Static Power vs. Temp.

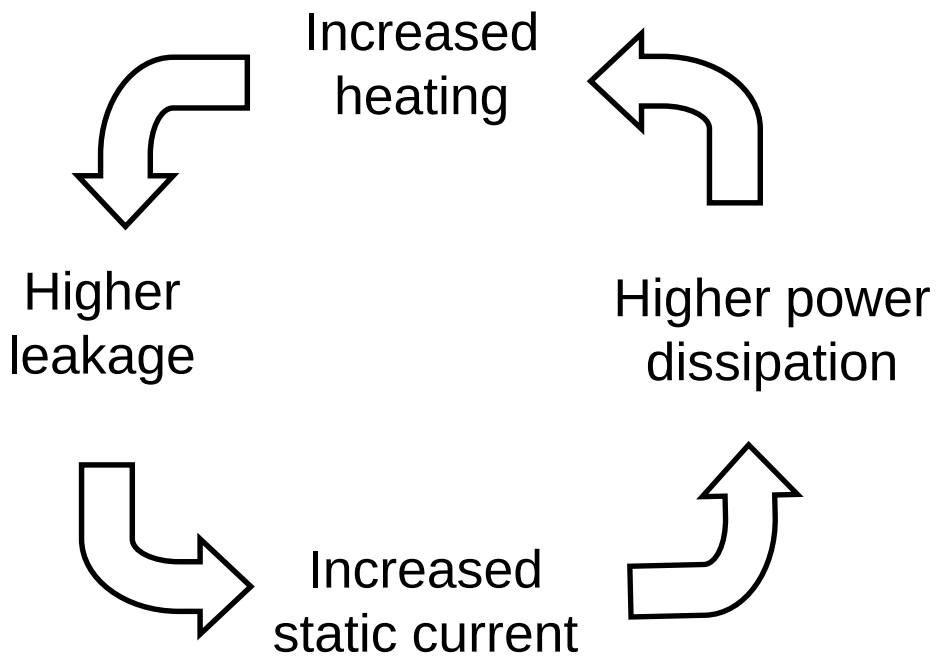


- Static power is more problematic in active mode for high performance microprocessors

source: synopsys

Thermal Runaway



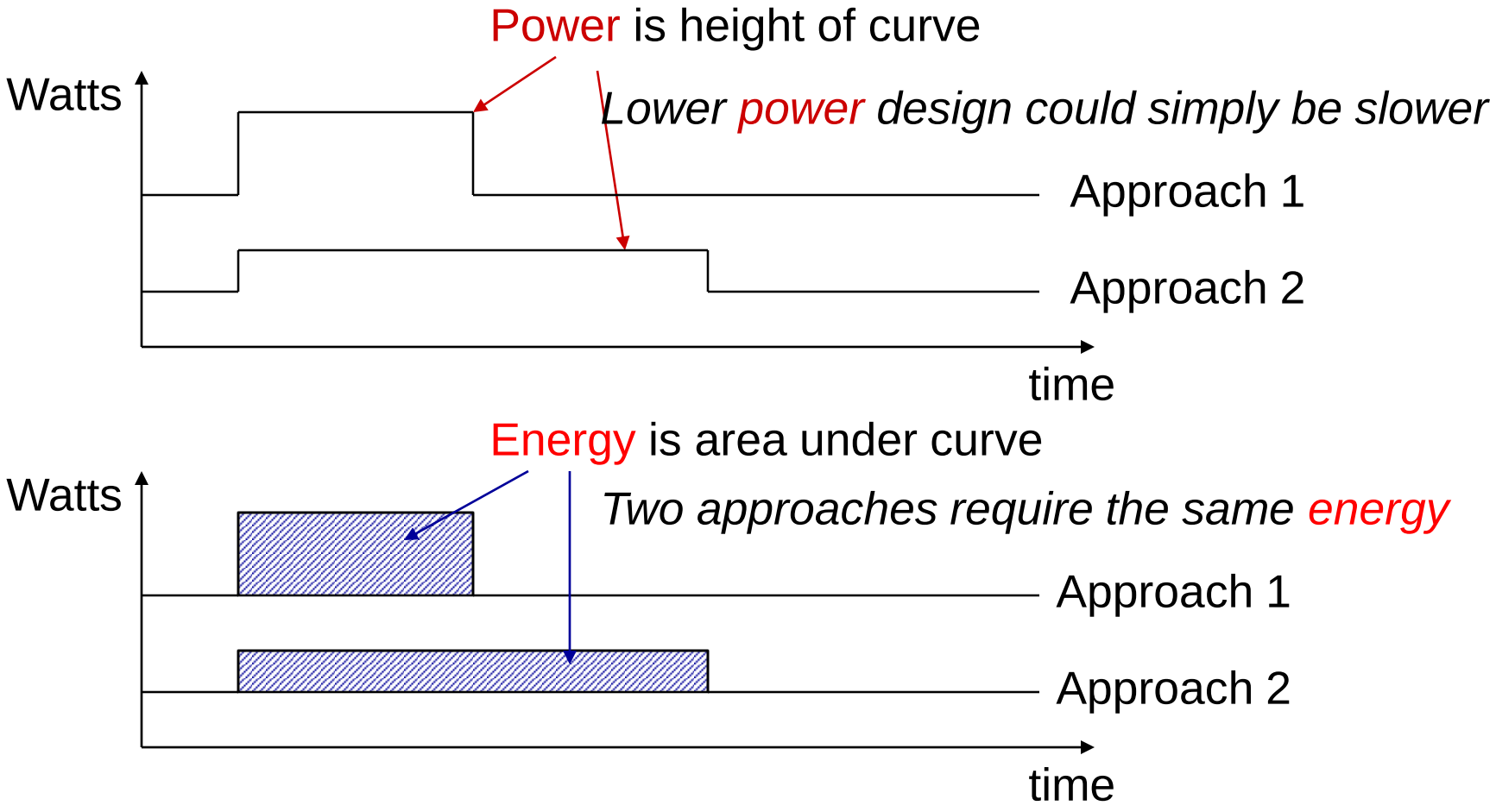


- Destructive positive feedback mechanism
- Leakage increases exponentially with temperature
- May destroy the test socket → thermal sensors required

source: synopsys

Borrowed from Xinfei Guo@JI, ECE4810J, 21FA

Power vs. Energy



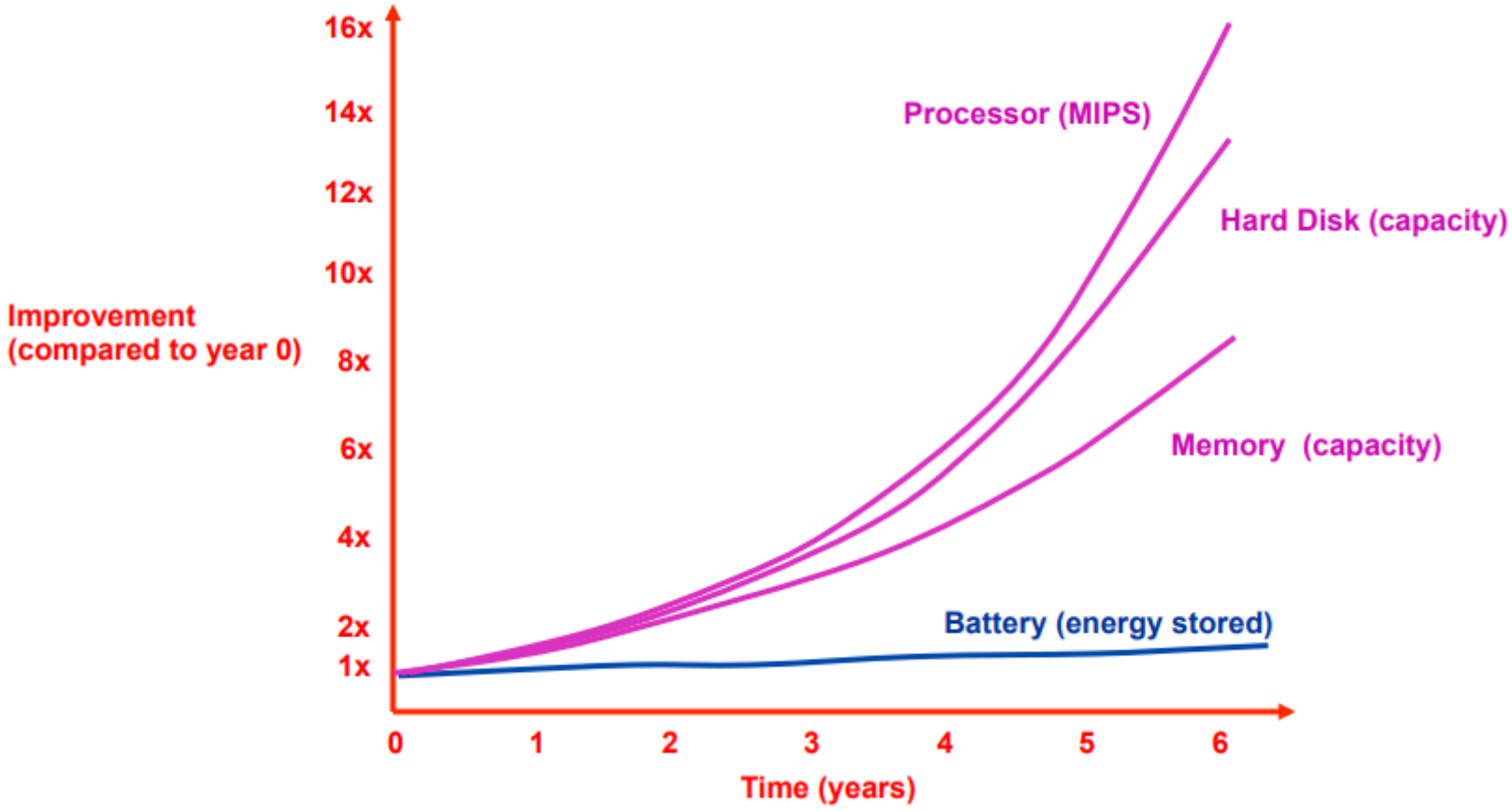
source: synopsys

Power and Energy

- Power consumption in Watts
  - Sets packaging limits
  - Indicates heat

- Energy efficiency in joules
  - Rate at which energy is consumed over Time
  - Determines battery life in hours
  - Energy = power * delay (joules = watts * seconds)
  - Lower energy number means less power to perform a computation at same frequency

Why worry about energy?



source: https://www.cse.wustl.edu/~lu/cse467s/slides/power.pdf

TDP – Thermal Design Power

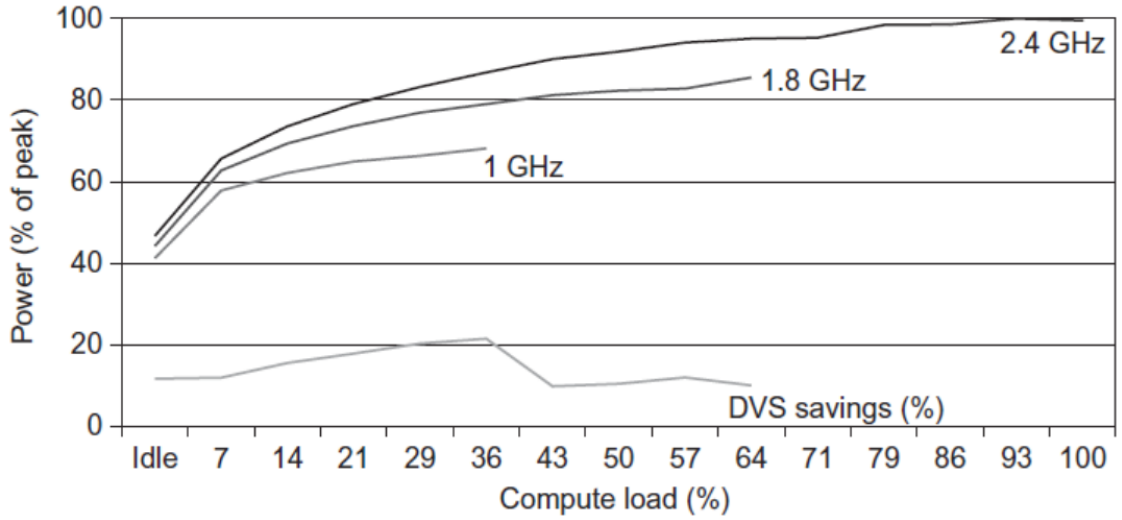
- A metric that is expressed in watts.
- It refers to the amount of power/heat a cooling system (i.e. heatsink, fan) is expected to dissipate to prevent overheating.
- The thermal design power is the maximum power a processor can draw for a thermally significant period while running commercially useful software.
- The TDP does not typically indicate the most power the chip could ever draw, but how much it draws under real load.
- Lower than peak power (1.5X higher), higher than average power consumption
- You can, generally, use TDP to gauge power consumption of a chip, in that the smaller the number for TDP, the lower power consumption by the CPU.

References

- https://www.intel.com/content/dam/doc/white-paper/resources-xeon-measuring-processor-power-paper.pdf

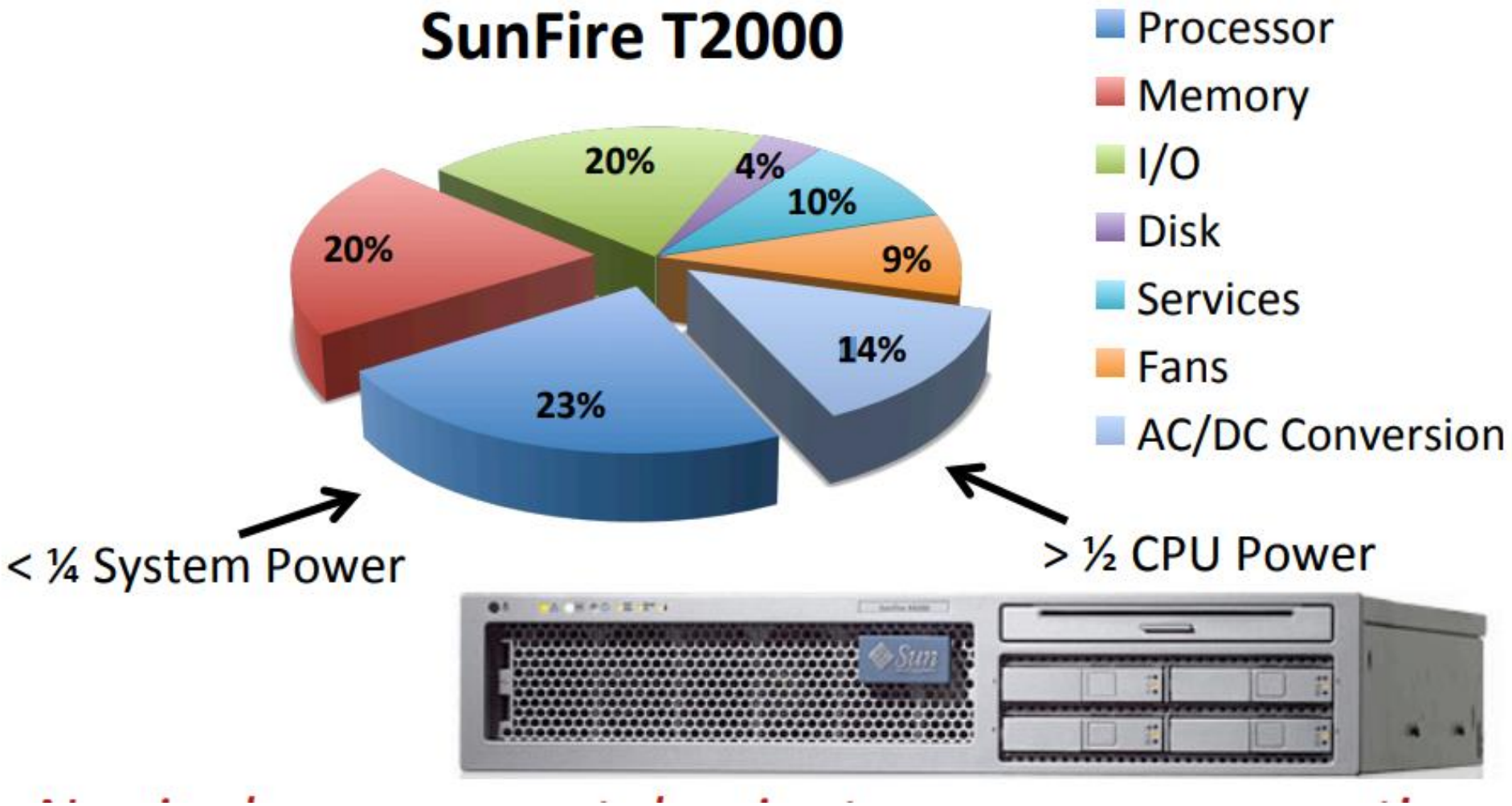
Power-Saving Features

- Voltage drops
  - Reduce power supply voltage.
- Dynamic Voltage-Frequency Scaling



- Low power state for DRAM, disks
- Overclocking, turning off cores - Disconnect parts from power supply when not in use

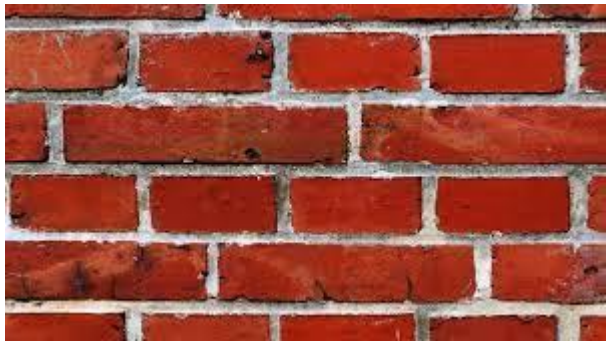
Server power breakdown

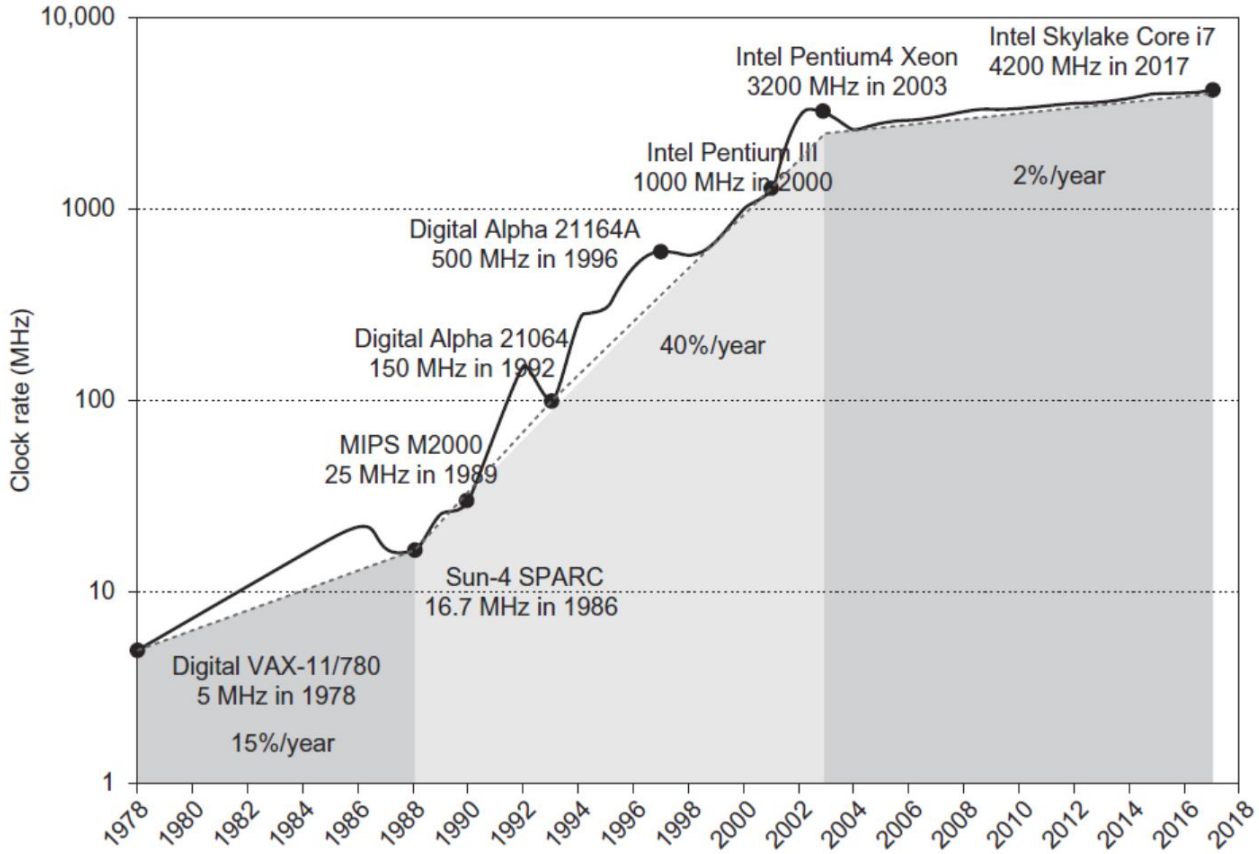


source: https://www.eecs.umich.edu/courses/eecs470/lectures/2.pdf

Power Wall

- Unfortunately, the increased the power dissipation of the CPU chip beyond the capacity of inexpensive cooling techniques.
- 3.3 GHz Intel Core i7 consumes 130W
- Heat must be dissipated from 1.5 x 1.5 cm chip
- This is the limit of what can be cooled by air
- We hit a wall!





https://thermtest.com/how-the-thermal-conductivity-of-clay-bricks-contributes-to-their-success-as-a-building-material

What is the solution?

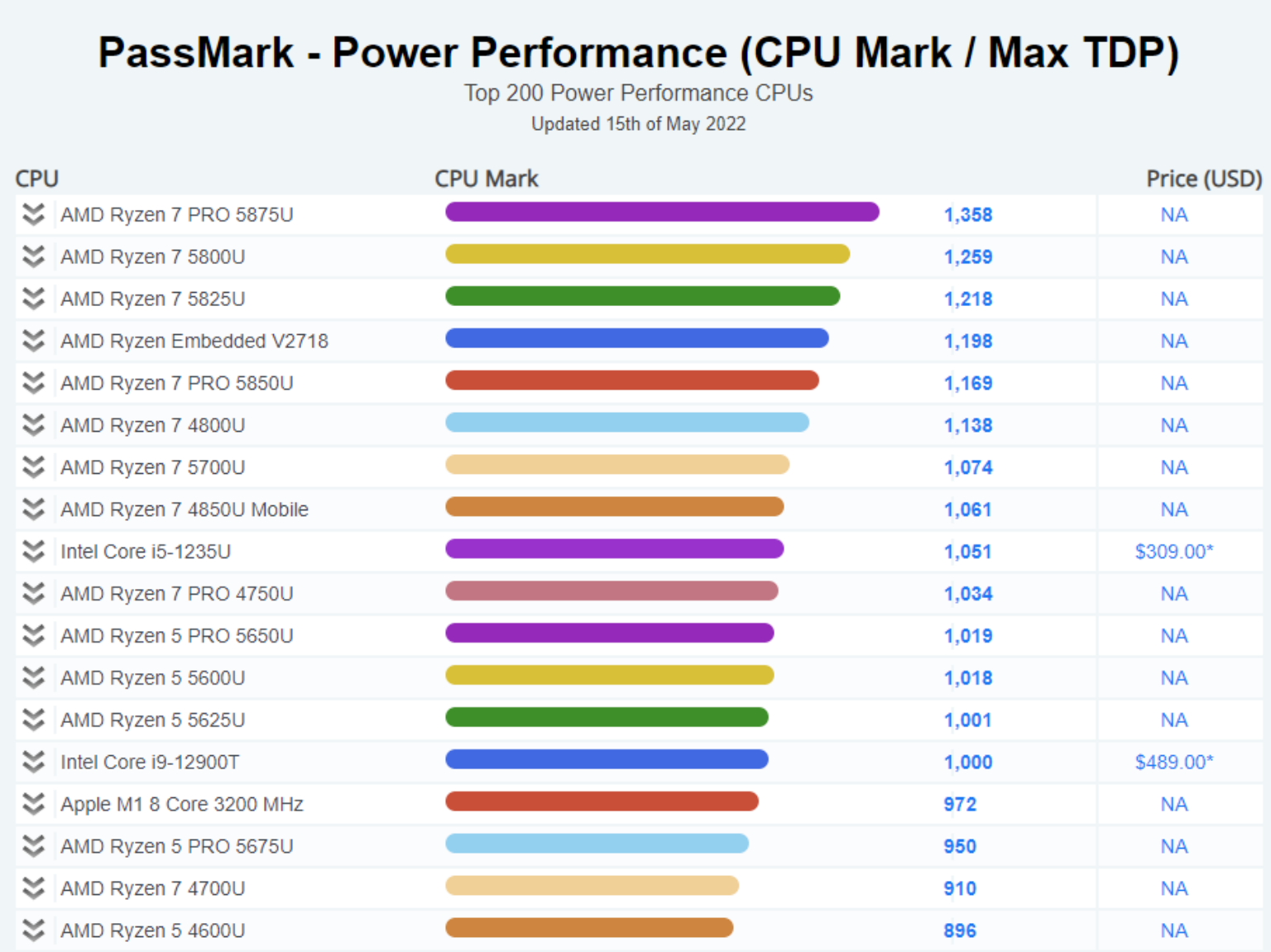
- The technique taken by IBM for implementation on their large z/10 and z/11 servers was to include sophisticated and costly cooling technologies. This allowed CPU clock rates in excess of 5.0 GHz. The commercial products today use 4.67 GHz CPU chips.
- The technique taken by commodity processor providers, such as Intel and AMD. They have moved to a multicore design, in which each CPU chip contains a moderate number of components. Due to cost constraints, the chips are cooled by simple fans and all–metal heat radiators.

source: http://www.edwardbosworth.com/My5155_Slides/Chapter01/ThePowerWall.htm

Liquid Cooling

source: Google

Power benchmark



PassMark - Power Performance (CPU Mark / Max TDP)

Top 200 Power Performance CPUs

Updated 15th of May 2022

CPU	CPU Mark	Price (USD)
AMD Ryzen 7 PRO 5875U	1,358	NA
AMD Ryzen 7 5800U	1,259	NA
AMD Ryzen 7 5825U	1,218	NA
AMD Ryzen Embedded V2718	1,198	NA
AMD Ryzen 7 PRO 5850U	1,169	NA
AMD Ryzen 7 4800U	1,138	NA
AMD Ryzen 7 5700U	1,074	NA
AMD Ryzen 7 4850U Mobile	1,061	NA
Intel Core i5-1235U	1,051	$309.00*
AMD Ryzen 7 PRO 4750U	1,034	NA
AMD Ryzen 5 PRO 5650U	1,019	NA
AMD Ryzen 5 5600U	1,018	NA
AMD Ryzen 5 5625U	1,001	NA
Intel Core i9-12900T	1,000	$489.00*
Apple M1 8 Core 3200 MHz	972	NA
AMD Ryzen 5 PRO 5675U	950	NA
AMD Ryzen 7 4700U	910	NA
AMD Ryzen 5 4600U	896	NA

https://www.cpubenchmark.net/power_performance.html

Combined Power-Performance Metrics

- Power-delay Product (PDP or Energy) = Pavg * t
  - PDP is the average energy consumed per switching event
- Energy-delay Product (EDP) = PDP * t
  - Takes into account that one can trade increased delay for lower energy/operation
- Energy-delay^2 Product (EDDP) = EDP * t
  - Why do we need so many formulas?!
  - We want a voltage-invariant efficiency metric! Why?
  - Power ~ ½ CV^2f, Performance ~ f (and V)

More Energy

- Energy per instruction (EPI): average amount of energy expended per instruction processed by the microprocessor
- MIPS/watt is the $\frac{\text{Joules}}{\text{Instruction}} = \frac{\frac{\text{Joules}}{\text{Second}}}{\frac{\text{Instructions}}{\text{Second}}} = \frac{\text{Watt}}{\text{IPS}}$ es: the rate at which the CPU is processing instructions, and the rate at which energy is being expended.
- MIPS/watt and EPI are ideal metrics for assessing power-efficiency in environments where **throughput performance** is the primary objective.
- higher throughput within a fixed power budget, lower EPI

More Energy

- It is important to note that MIPS/watt and EPI do not consider the amount of time (latency) needed to process an instruction from start to finish.
- Other metrics such as $MIPS^2$ /watt (related to ED) and $MIPS^3$ /watt (related to ED^2) assign increasing importance to the time required to process instructions, and are thus used in environments in which latency performance is the primary objective.

An example

- Currently have a processor design:
  - 80W, 1 BIPS, 1.5V, 1GHz
  - Want to reduce power, willing to lose some performance
- Cache Optimization:
  - IPC decreases by 10%, reduces power by 20% => Final Processor: 900 MIPS, 64W
  - Relative E = MIPS/W (higher is better) = 14/12.5 = 1.125x
  - Energy is better, but is this a “better” processor?

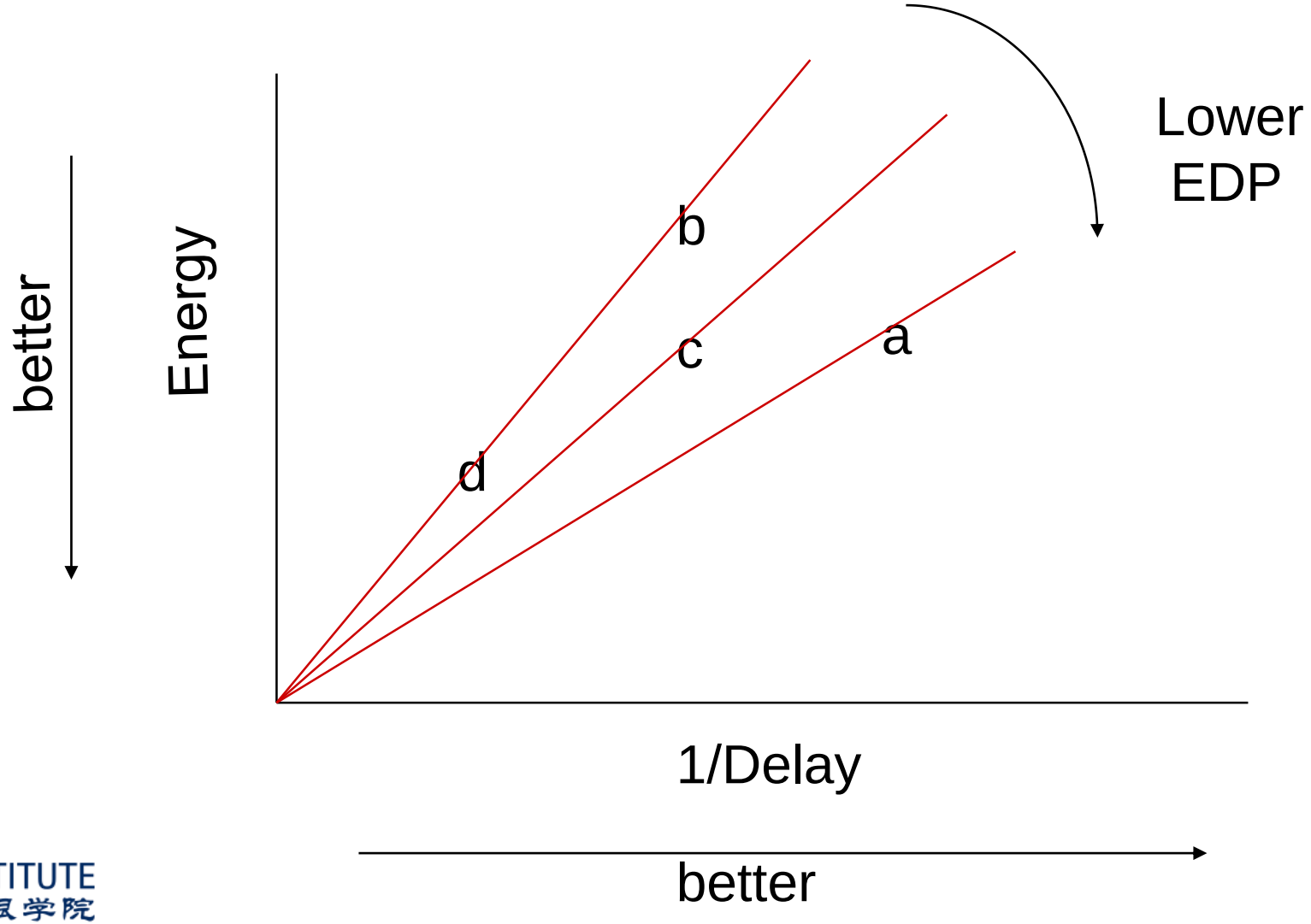
source: Prof. David Brooks@Harvard

Not necessarily

- 80W, 1 BIPS, 1.5V, 1GHz
- Cache Optimization:
  - IPC decreases by 10%, reduces power by 20% => Final Processor: 900 MIPS, 64W
  - Relative E = MIPS/W (higher is better) = 14/12.5 = 1.125x
  - Relative EDP = $MIPS^2/W$ = 1.01x
  - Relative ED^2P = $MIPS^3/W$ = .911x
- What if we just adjust frequency/voltage on processor?
  - How to reduce power by 20%?
  - $P = CV^2F = CV^3$ => Drop voltage by 7% (and also Freq) => .93*.93*.93 = .8x
  - So for equal power (64W)
    - Cache Optimization = 900MIPS
    - Simple Voltage/Frequency Scaling = 930MIPS

Understanding Tradeoffs

- Which design is the “best” (fastest, coolest, both) ?



source: synopsys

Power vs. Computer Architecture

- General-purpose vs. Application specific

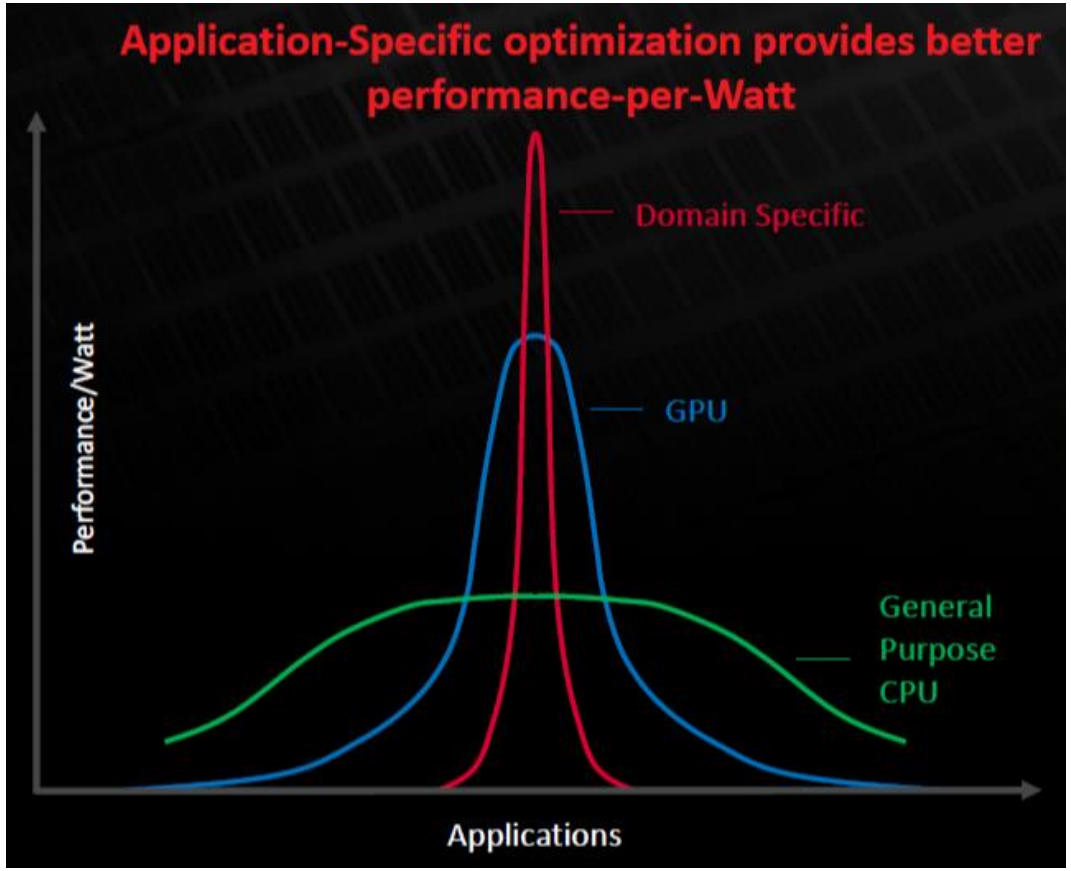
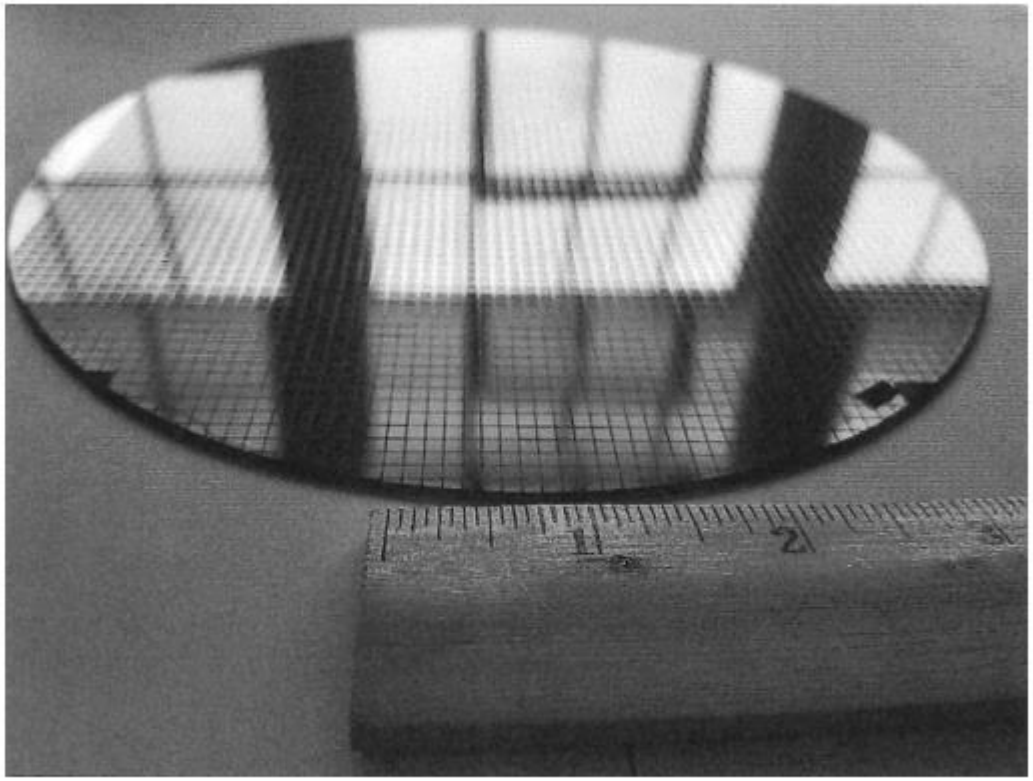


Figure: AMD

Die Area and Cost

- There are significant side effects that die area has on the fixed and other variable costs
- Die sizes of a processor is usually about 10-15 mm on a side.
- The die is produced in bulk from a larger wafer, perhaps 30 cm in diameter.
- Silicon wafers and processing technologies are not perfect. Defects randomly occur over wafer surface



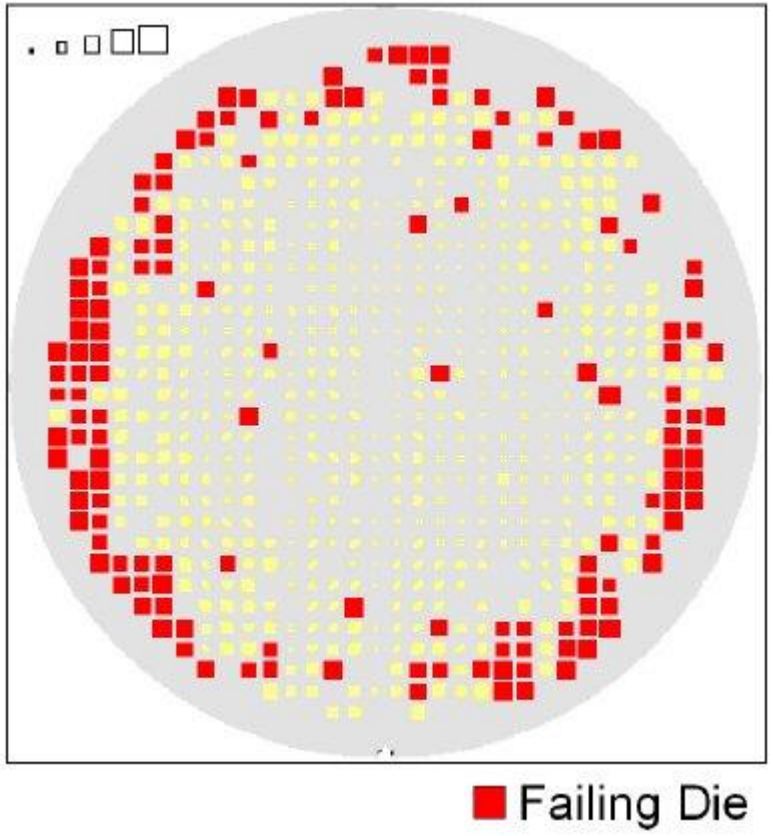
A RISC-V Wafer

Trends in Cost

- Cost driven down by learning curve
  - Yield
- DRAM: price closely tracks cost
- Microprocessors: price depends on volume
  - 10% less for each doubling of volume

Defects

- silicon and technology processes are imperfect.
- defect can lead to failing dies



source: https://www.researchgate.net/publication/3953891_Statistical_post-processing_at_wafersort-an_alternative_to_burn-in_and_a_manufacturable_solution_to_test_limit_setting_for_sub-micron_technologies/figures?lo=1

Integrated Circuit Cost

- Integrated circuit

$$\text{Cost of integrated circuit} = \frac{\text{Cost of die} + \text{Cost of testing die} + \text{Cost of packaging and final test}}{\text{Final test yield}}$$

$$\text{Cost of die} = \frac{\text{Cost of wafer}}{\text{Dies per wafer} \times \text{Die yield}}$$

$$\text{Dies per wafer} = \frac{\pi \times (\text{Wafer diameter}/2)^2}{\text{Die area}} - \frac{\pi \times \text{Wafer diameter}}{\sqrt{2 \times \text{Die area}}}$$

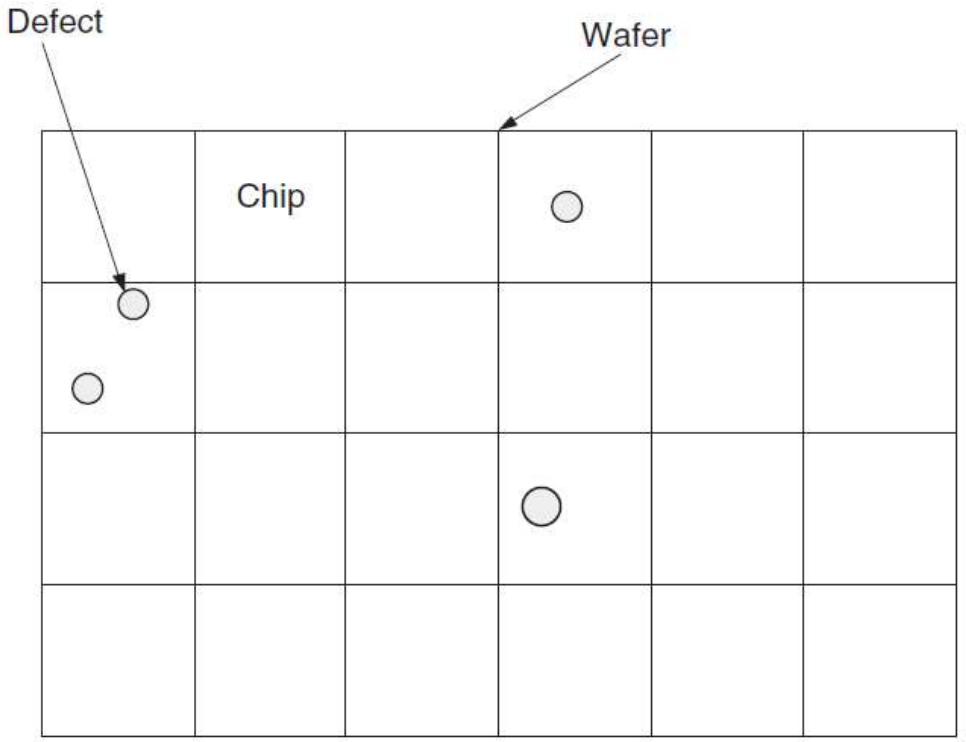
- Bose-Einstein formula:

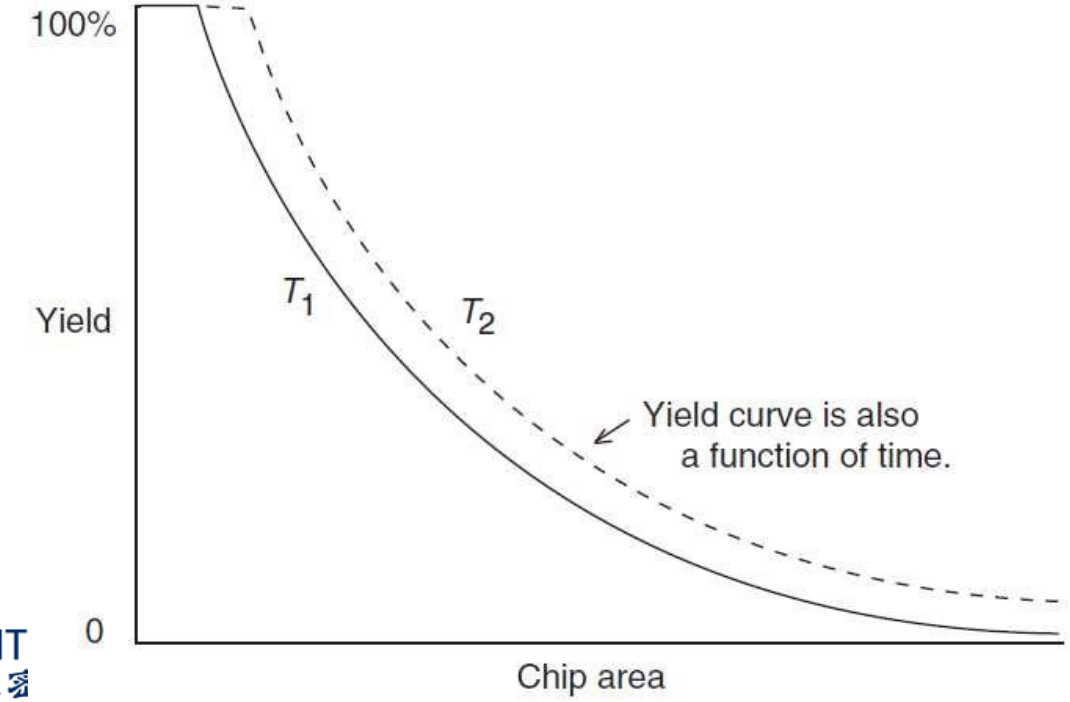
$$\text{Die yield} = \text{Wafer yield} \times 1/(1 + \text{Defects per unit area} \times \text{Die area})^N$$

- Defects per unit area = 0.016-0.057 defects per square cm (2010)
- N = process-complexity factor = 11.5-15.5 (40 nm, 2010)

Wafer Defects

- Defects randomly occur over the wafer surface.
- Large SoC chip area requires an absence of defects over that area





Reliability

- Reliability is a measure of functionality over time. If a processor completely stops working, it is deemed unreliable. But in many cases performance degrades over a period of years, whether that's due to electromigration, layer upon layer of software patches, memory bit failures, or a host of other issues that can crop up. Reliability in these cases may be more subjective than definitive.
- Reliability is also known as Dependability and Fault-Tolerance

source: https://semiengineering.com/are-chips-getting-more-reliable/

Reliability is important

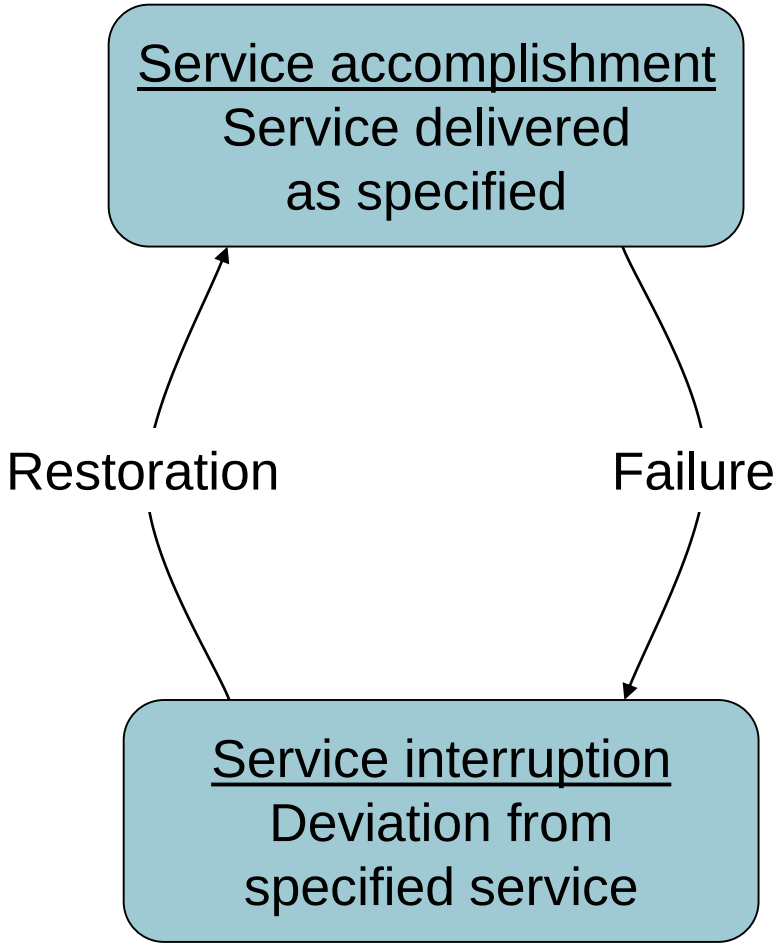
- In 2021, Tesla recalled over 15800 cars due to Flash Memory Failure
- NAND flash wears out in an eMMC device (used in massive touchscreen displays)
- Reliability is about lifetime!

Figure: https://www.tomshardware.com/news/flash-memory-wear-killing-older-teslas-due-to-excessive-data-logging-report

Dependability

- Processors not only need to be fast; they need to be dependable.
- **Dependability** is the ability to provide services that can be trusted within a time-period.
- Example: Dependability via Redundancy

Dependability



- Fault: failure of a component
  - May or may not lead to system failure

Dependability Measures

- Reliability: mean time to failure (MTTF)
- Service interruption: mean time to repair (MTTR)
- Mean time between failures
  - MTBF = MTTF + MTTR
- Availability = MTTF / (MTTF + MTTR)
- Improving Availability
  - Increase MTTF: fault avoidance, fault tolerance, fault forecasting
  - Reduce MTTR: improved tools and processes for diagnosis and repair

Factors

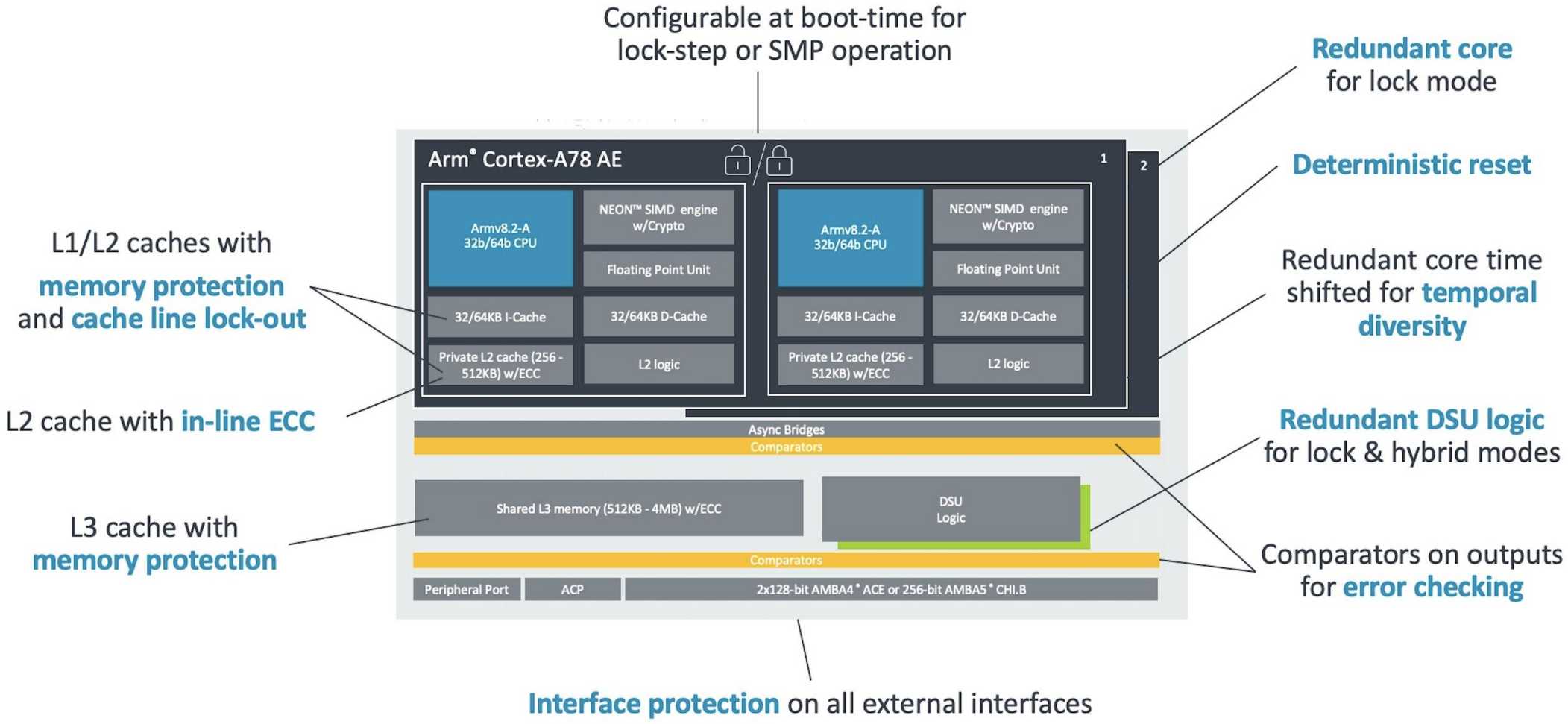
- Reliability is related to die area, clock frequency, and power.
- Die area increases the amount of circuitry and the probability of a fault.
- It also allows the use of error correction and detection techniques.
- Higher clock frequencies increase electrical noise and noise sensitivity.
- Faster circuits are smaller and more susceptible to radiation.

RAID: Redundant Arrays of (Inexpensive) Disks

- Data is stored across multiple disks
- Files are “striped” across multiple disks
- Redundancy yields high data availability
- Availability: service still provided to user, even if some components failed

Image: http://marketexplorereports.blogspot.com/2018/01/global-redundant-array-of-independent.html

Redundancy can also help security

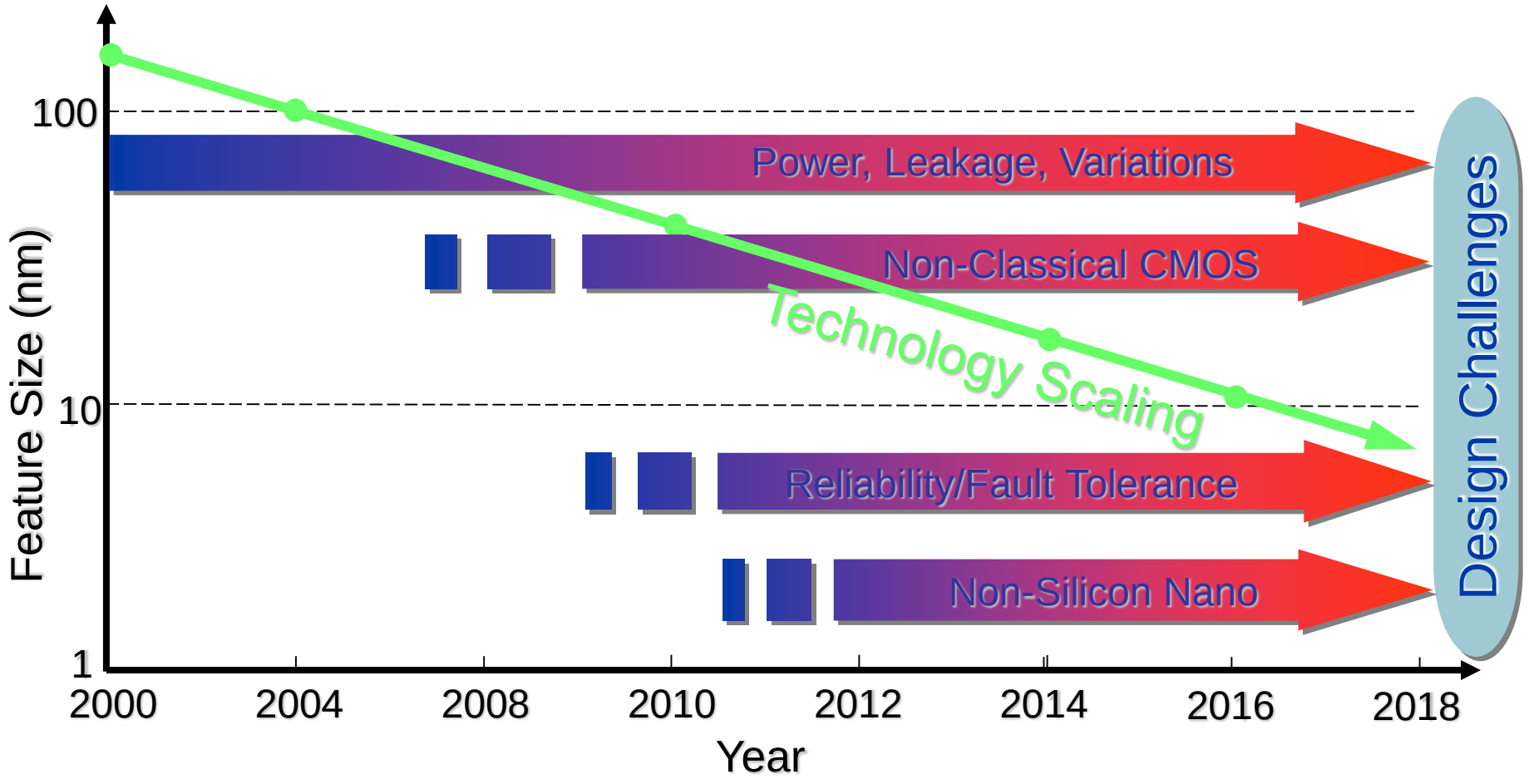


source: https://community.arm.com/arm-community-blogs/b/embedded-blog/posts/arm-cortex-a78ae-on-the-road-to-an-autonomous-future

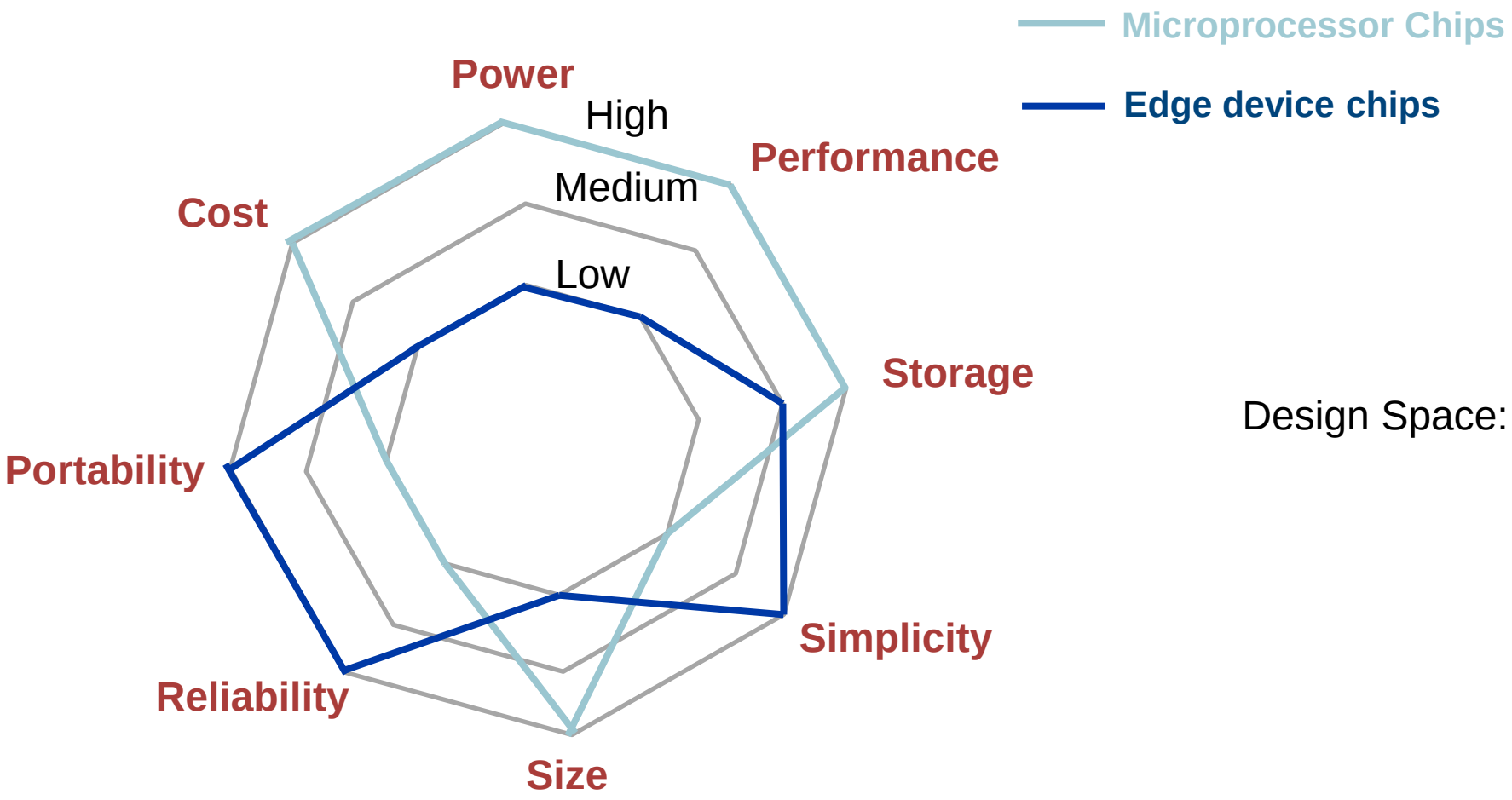
Put it all together

- What does “Quantitative” mean here?
  - Defining quantitative ways to measure how good an architecture can be.
- Why do we need to know these metrics?
  - Computer architectures eventually need to meet requirements from all these metrics.
  - Designing an optimal architecture need to trade one with another.

Summary: Design Challenges



Design Space is massive!



Design Space:

Adapted from [R. Sanchez-Iborra, etc. IEEE Circuits and Systems Magazine, 2020]

Where are we Heading?

- T3: Fundamental Processors

Acknowledgement

Slides in this topic are inspired in part by material developed and copyright by:

- ARM Courseware
- Prof. Ron Dreslinski @ UMich, EECS 470
- Prof. David Brooks @ Harvard
- Xinfei Guo@JI, VE481 2021 FA

Action Items

- Check the lab schedules and get ready
- HW#1 is upcoming
- Reading Materials
  - Ch. 1.5 - 1.8, 1.10