POLITECNICO DI MILANO

HEAP LAB





Thermal management in computer architectures
Embedded systems – Prof. William Fornaciari

Federico Terraneo
federico.terraneo@polimi.it

Outline

The thermal problem

- Power consumption of digital logic
- Effects of temperature on integrated circuits

Modeling thermal phenomena

- Thermal simulators
- IR thermography
- Thermal test chips

Heat dissipation

- Natural convection
- Heat sink and fans
- Advanced cooling solutions

Thermal control policies

- With on/off actuators
- With DVFS

The thermal problem

CMOS logic draws power through two main mechanisms

- Active power

$$P_{active} = C\, V_{dd}^{2} f$$

  - Increases linearly with frequency and quadratically with the supply voltage
  - Parameters cannot be changed arbitrarily, to operate at a given frequency, a minimum voltage is required

- Leakage power

$$P_{leak} = I_0\, v_T^{2}\, e\hat{}((v_{gs} - v_{th} - v_{off})/(\eta v_T))(1-e\hat{}(-v_{ds}/v_T))$$ [1]

Formula is more complex but in a nutshell:

  - Increases exponentially with temperature
  - Increases as transistor size decreases, worsens as Moore's law progresses

[1] https://arxiv.org/pdf/1809.03147.pdf

The thermal problem

Power consumed by an integrated circuit is released as heat that increases the chip temperature

High temperatures have negative effects on ICs

- Increase gate delay, slows down maximum operating frequency
  - Requires to increase V_{dd} to maintain performance level
    - → Causing additional power consumption
- Lower MTBF
  - Failure modes of integrated circuits are accelerated by high temperatures
    - → Electromigration
    - → NBTI
- Thermal runaway
  - Leakage power increases exponentially with temperature, causing further temperature increase
  - High temperatures can initiate a positive feedback loop leading to IC destruction

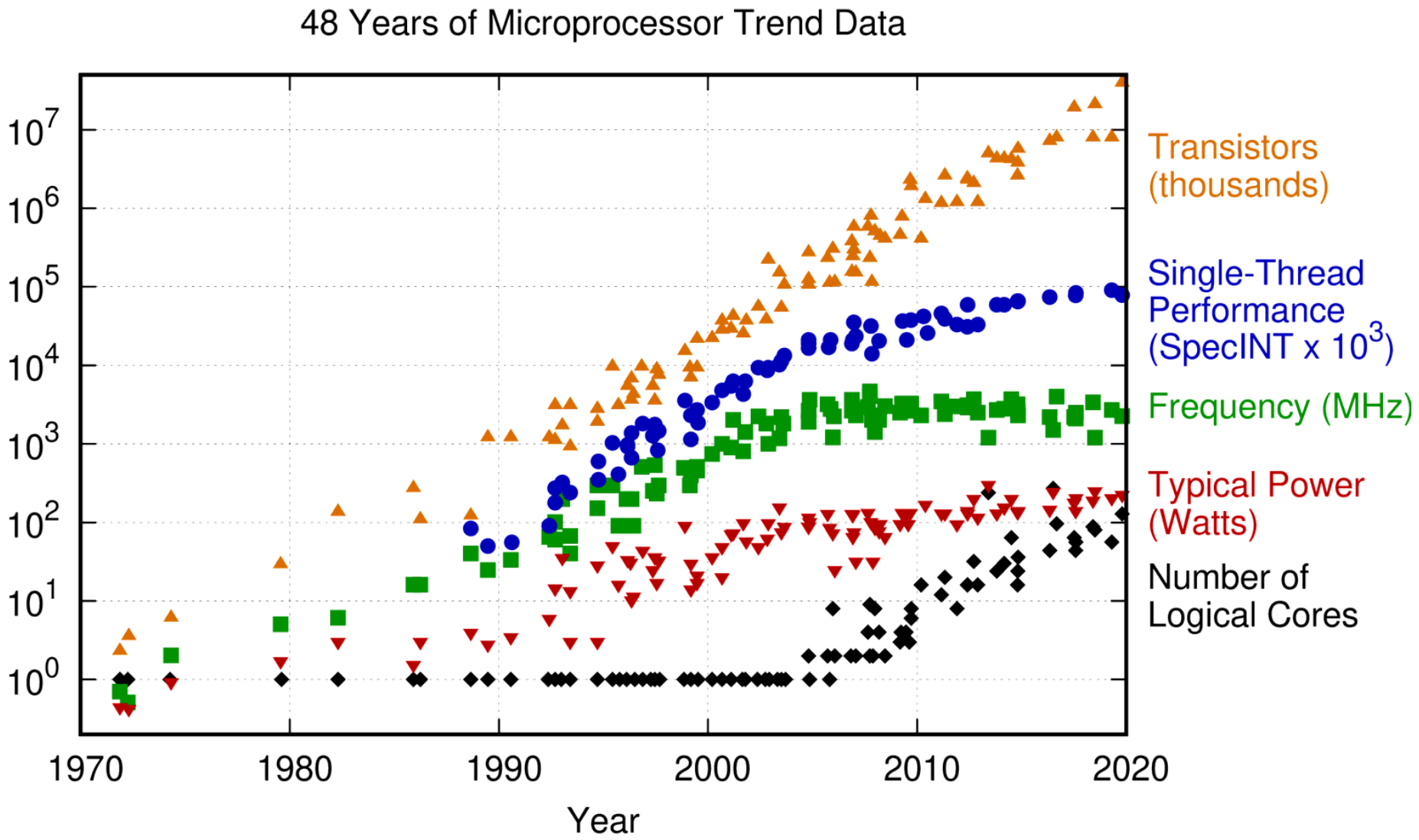
The thermal problem

Limitations in the power that can be dissipated from a chip shaped the evolution of computing systems

History

- Mid 1990s: Introduction of forced convection fan cooling (Pentium)
- Early 2000s: CPU clock frequency stops increasing (power/thermal wall) propelling the multicore revolution
- 2010s: CPUs can no longer operate indefinitely at their maximum clock frequency (Intel Turbo Boost introduced)
- Mid 2010s: high performance integrated circuits can no longer afford to power on all their transistors without thermal damage (Dark silicon)

The thermal problem



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 2010-2019 by K. Rupp

Modeling thermal phenomena

Heat conduction in solids is modeled using 3D Fourier equation

$$\frac{\partial T(x,y,z,t)}{\partial t} = \frac{\lambda(x,y,z)\nabla^2 T(x,y,z,t) + Q(x,y,z,t)}{c(x,y,z)\rho(x,y,z)}$$

- Partial differential equation
  - T is the chip temperature
  - Q is the power dissipated inside the chip [W/m^2]
  - ρ, c, λ, are the density, specific heat and thermal conductivity
  - ∇^2 is the Laplacian operator
- To be solved, a spatial discretization method is required
  - Finite volumes
  - Finite elements

Modeling thermal phenomena

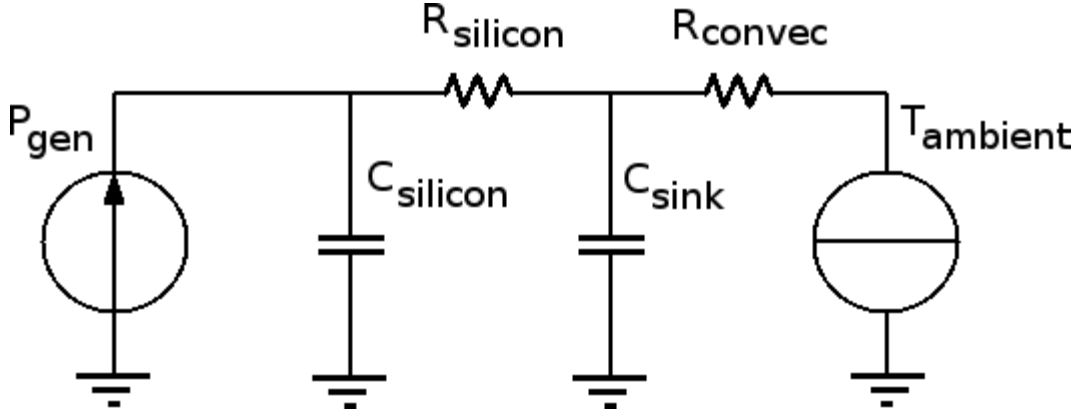
Electrical equivalent of thermal systems

- Simplified analysis method

- Thermal conductor → Electrical conductance
- Thermal capacity → Electrical capacitor
- Thermal power source → Electrical current source
- Fixed temperature → Voltage source

Modeling thermal phenomena

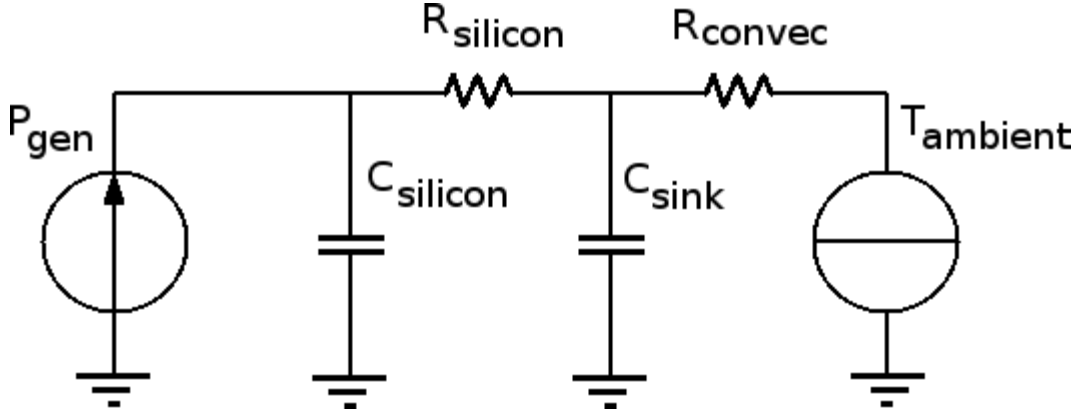
A simplistic chip thermal model



- P_{gen} is the power generated inside the chip
- $C_{silicon}$ is the silicon thermal capacity
- $R_{silicon}$ is the resistance between the chip and the heat sink
  - silicon bulk, heat spreader, thermal paste, etc.
- C_{sink} the heat sink thermal capacity
- R_{convec} models heat convection towards the ambient

Modeling thermal phenomena

A simplistic chip thermal model



Important considerations

- $C_{silicon}$ is very small
  - Chips are small and thin
- $R_{silicon}$ is non-negligible

→ The problem exhibits **Fast thermal dynamics**

- Thermal phenomena are usually considered slow, but this is not the case for the silicon temperature
  - Temperature changes of 10..20°C in <100ms are common

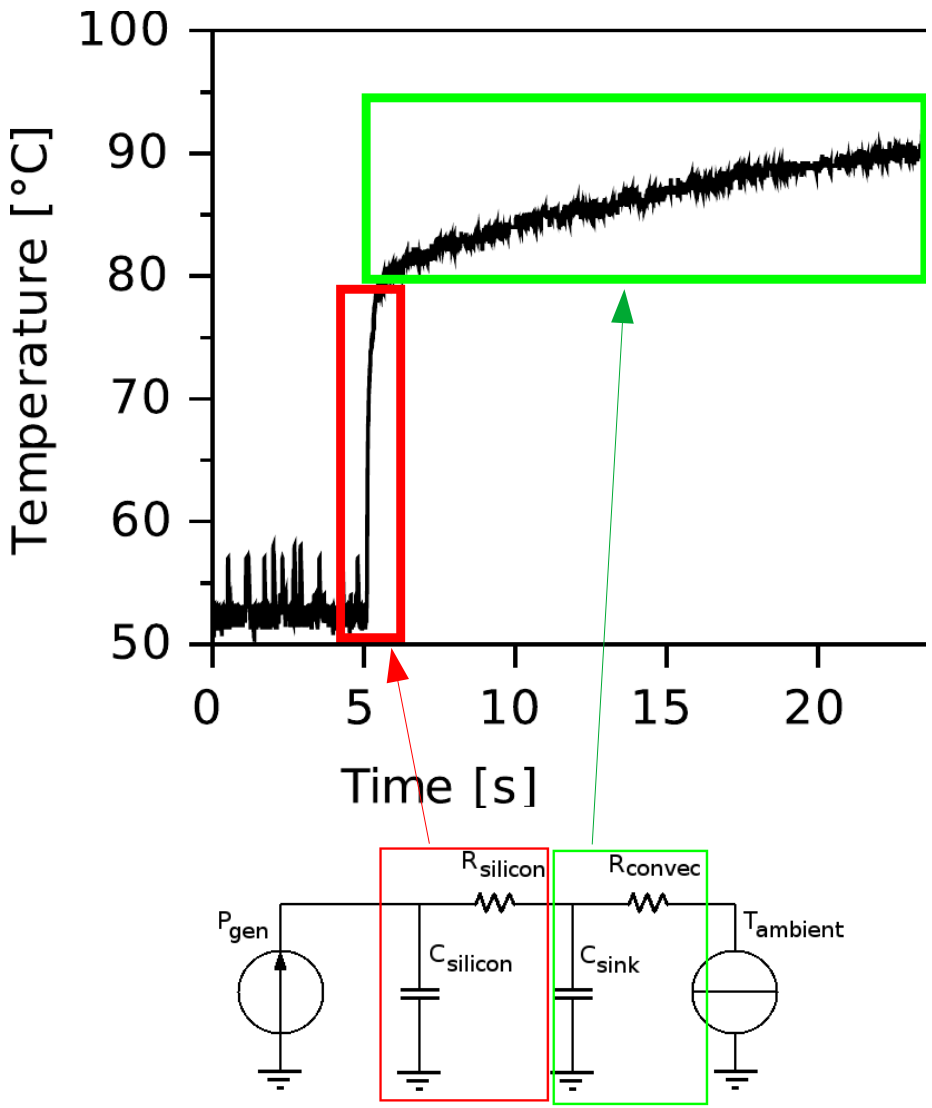
Modeling thermal phenomena

Experimental setup

- Intel Core-i5 6600K
- Instrumented Linux kernel for sampling CPU temperature at 1KHz
- cpuburn program run @ t=5s

Thermal dynamics

- Baseline temperature 54°C (noisy)
- Initial temperature rise caused by fast thermal dynamics
  - 70°C reached in 50ms
  - 80°C reached in 600ms
- Subsequent temperature increase happens due to the heat sink thermal dynamics
- 90°C reached in >20s



https://link.springer.com/chapter/10.1007/978-3-319-91962-1_5

Modeling thermal phenomena

Thermal modeling is required for the design of high-performance chips

- The simplisitic model is too simple to design chips
  - Real thermal simulators produce a 2D map of the chip temperatures
  - Needed to optimize the chip floorplan
- Despite its simplicity, the simplistic model provides useful insights
  - Can explain the temperature trace of a real microprocessor
  - Useful for teaching the main phenomena involved

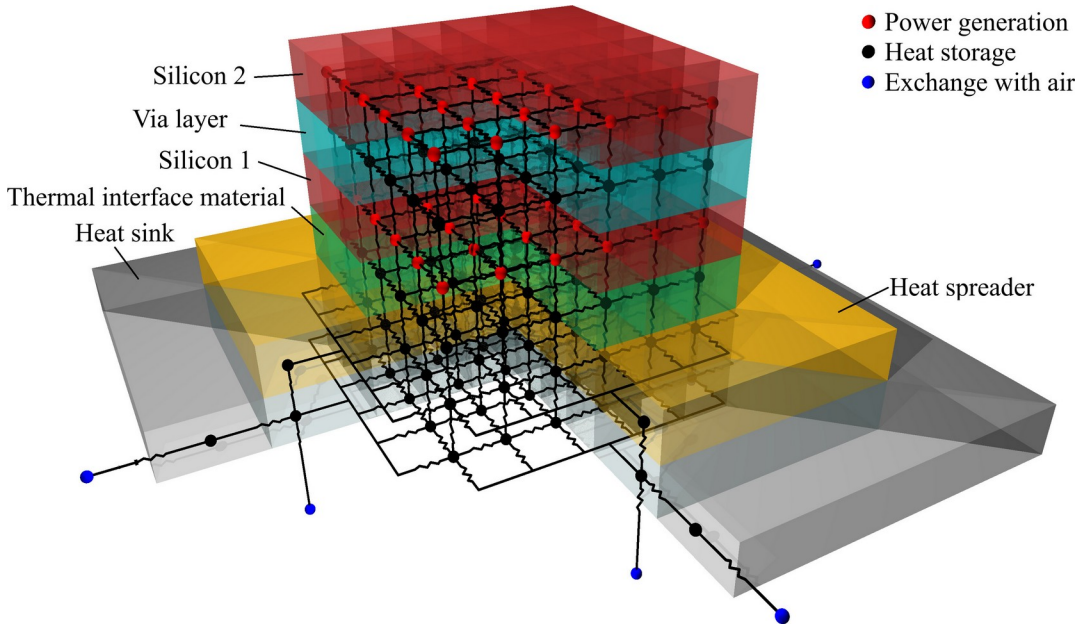
Modeling thermal phenomena

Thermal simulators

- Integrate the differential equations of heat transfer
  - Most common approach: finite volumes
    - → Divide the model to be simulated in a large number of small volumes
    - → Assume temperature is uniform inside each volume
    - → Compute the heat transfer across volumes

Structure of a typical simulator

- Inputs
  - Chip geometry
  - Material properties
  - Power trace
- Simulation engine
  - A differential equation solver
- Output
  - Temperature trace



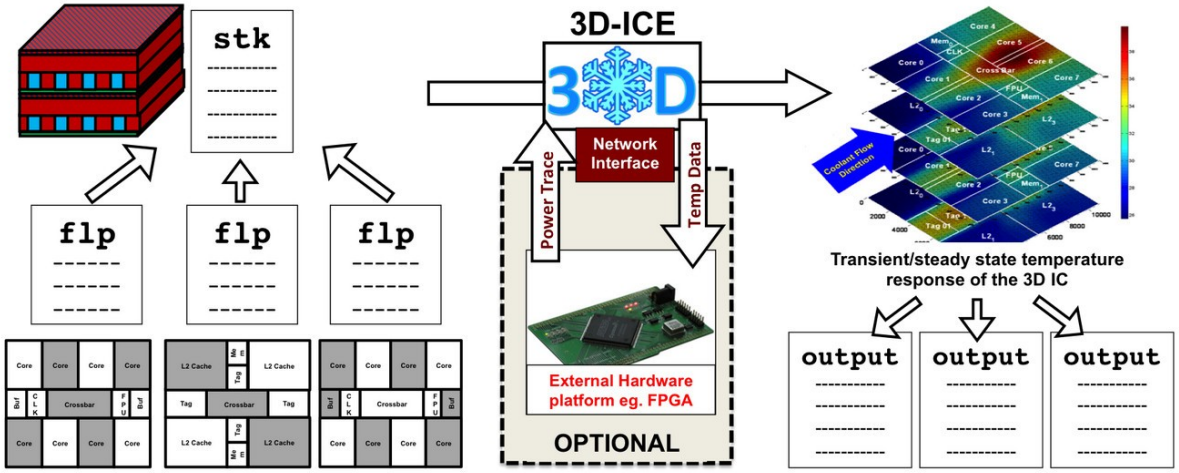
Modeling thermal phenomena

3D-ICE is an example of a thermal simulator

3D-ICE stands for **3D Interlayer Cooling Emulator**. It is a Linux based Thermal Emulator Library written in C, which can perform transient thermal analyses of vertically stacked 3D integrated circuits with inter-tier Microchannel Liquid Cooling. It is based on Compact Transient Thermal Modeling (CTTM) of solids and liquids.

Click here full list of publications and documentation related to 3D-ICE.

Also, check out our new FAQ page.

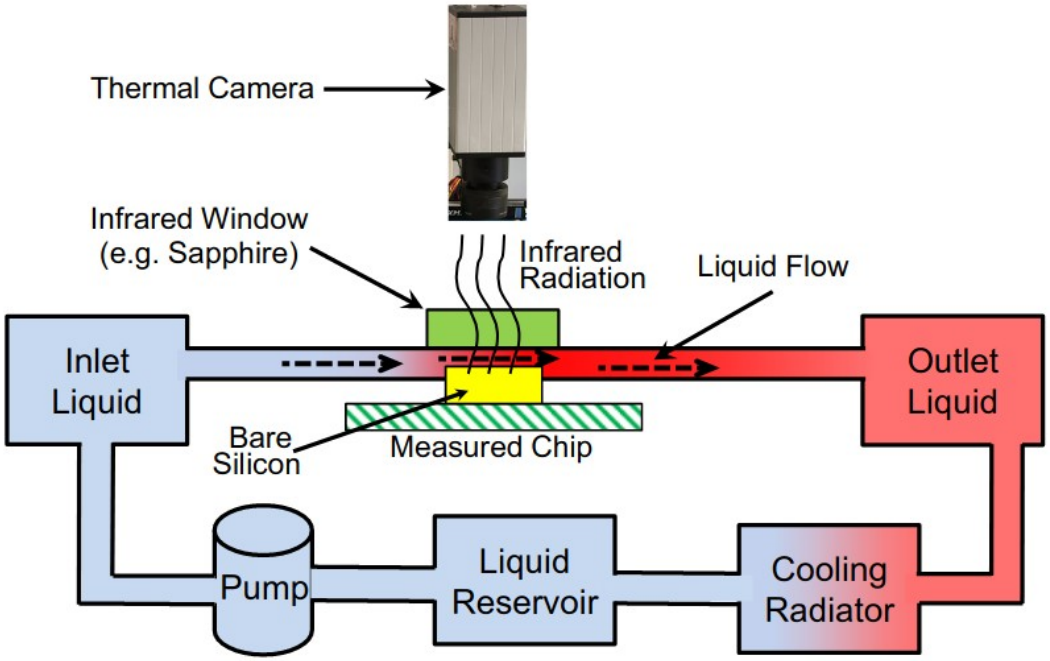


https://www.epfl.ch/labs/esl/research/open-source-software-projects/3d-ice/

Modeling thermal phenomena

IR thermography

- A way of directly observing the temperature map of an IC
  - The IC under test is decapped to expose the silicon
  - The IC is immersed in a circulating bath with an oil transparent to IR
  - An IR thermal camera is used to directly observe the temperature map
- Adavantages
  - High resolution thermal maps of real chips
- Disadvantages
  - The silicon temperatures of this setup are different compared to the normal operating conditions

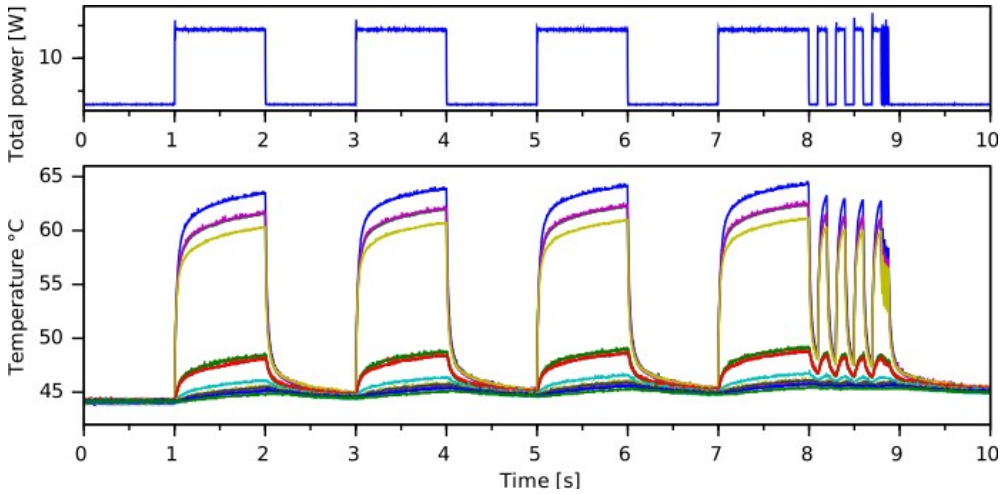


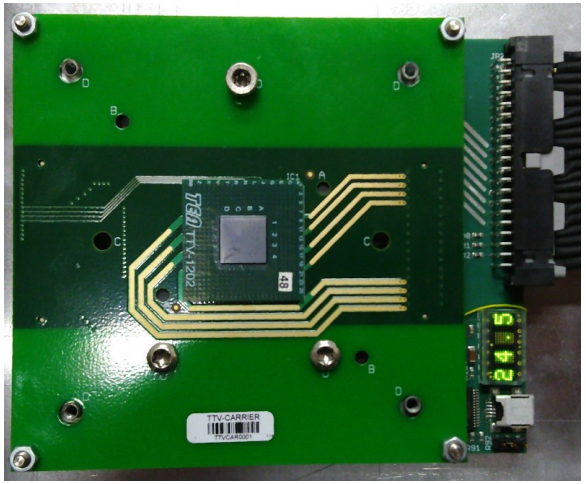
https://ces.itec.kit.edu/img/Lucid_Infrared_Thermography_of_Thermally_Constrained_Processors.pdf

Modeling thermal phenomena

Thermal test chips (TTCs)

- A TTC is a custom integrated circuit with
  - An array of power sources
  - An array of temperature sensors
- Provide a way to measure silicon temperatures and their spatial distribution
  - Can connect an arbitrary heat sink type
    - → Unlike IR thermography which forces the use of an oil bath
  - Useful for designing thermal models of heat sinks





https://link.springer.com/chapter/10.1007%2F978-3-030-27562-4_13

Heat dissipation

Heat produced inside an IC must be dissipated towards the ambient

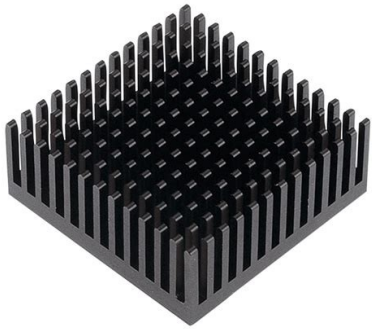
- Improvements in heat sinking allow to increase maximum dissipation
  - More performance for a given semiconductor manufacturing process
  - Subject of ongoing research
- Different heat sinking solutions
  - Passive cooling
  - Forced air cooling (fans)
  - Next generation cooling solutions
- Need to consider cost-effectiveness
  - Performance / cost ratio
  - Thermal control policies can help mitigate dissipation peaks

Heat dissipation

Passive cooling

- IC package exchanges heat directly with the ambient
  - Primary heat transfer path: natural convection
    - → A slow air movement induced by density difference between cold and hot air
  - Only usable for low power ICs, up to a few Watts

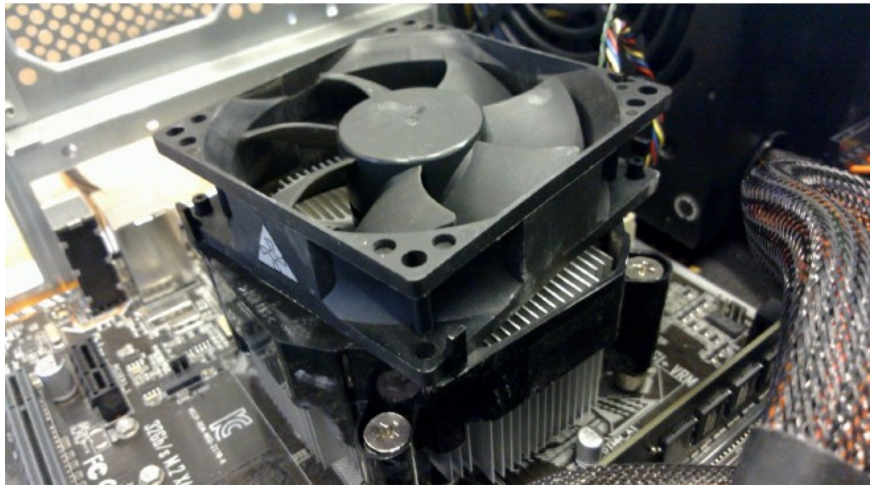
- Heat sinks
  - Components made of a material with high thermal conductivity (metal)
    - → Increases surface area where heat can be transferred to the ambient
    - → Geometry optimized to maximize surface area minimizing amount of material
  - Only usable for powers up to a few tens of Watts
    - → Size becomes prohibitively large when relying on natural convection alone



Heat dissipation

Forced air cooling

- The dominant heat sinking solution for high-performance ICs
  - In the desktop domain
  - In the datacenter/cloud domain
- Use fans to force air flow
  - Greatly increases heat transfer coefficient compared to natural convection
- Allow to use reasonably sized heat sinks
- Usable for powers up to a few hundred of Watts



Heat dissipation

Next generation cooling solutions

- Water cooling
  - Commercially available solution, gaining popularity
  - Replaces air with water to further increase heat transfer coefficient
  - Still requires a radiator to cool the water
- Evaporative cooling
  - Takes advantage of latent heat of vaporization uses special fluids with suitable boiling points
  - As long as the fluid is two-phase temperature cannot increase
- Thermoelectric cooling
  - Uses Peltier effect to pump heat from the IC
  - Useful for removing heat from localized hot spots

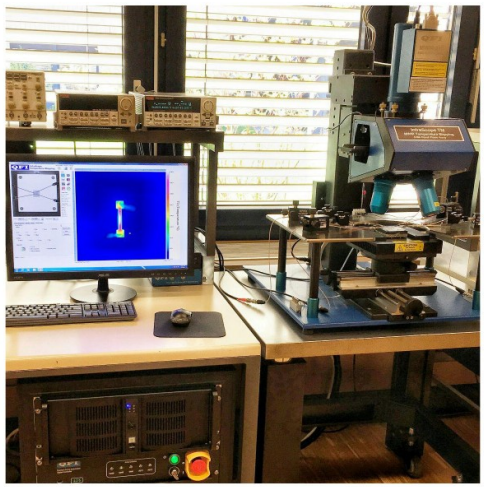


Fig. 5. Infra-Red microscope used for obtaining MPSoC temperature maps

Thermal control policies

Heat that can be dissipated from an MPSoC is limited

- Even more so when considering heat sink cost-effectiveness
- Dark silicon: current ICs heat dissipation exceed current heat sink technology

However, power consumed by an MPSoC is not constant

- Varies with CPU load, cache miss pattern, …

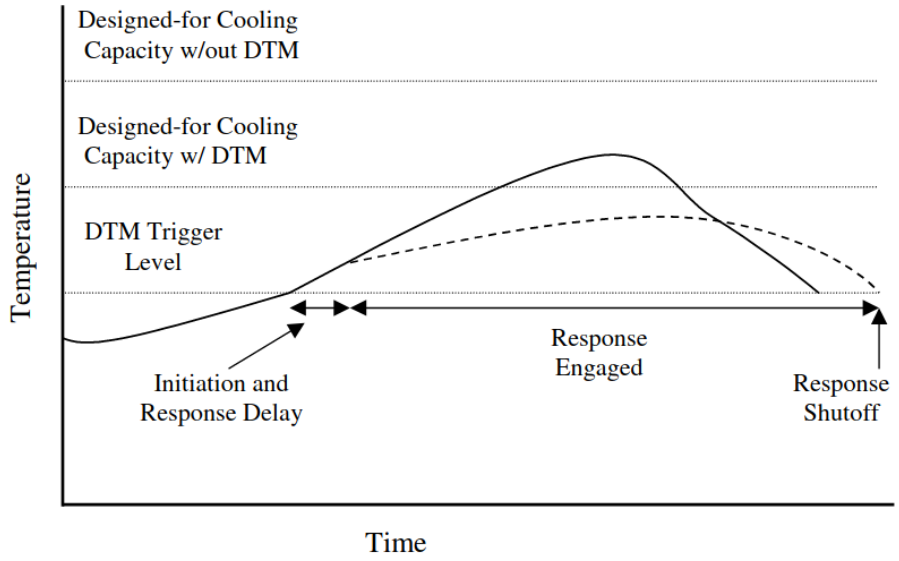
Idea: design an MPSoC that, in the worst case, draws more power than can be dissipated

- And use a thermal control policy to limit power during peaks
- Unlike heat sink improvements, the policy does cause a performance reduction during peaks
- But on average, the MPSoC can be faster than if it were limited to worst-case heat dissipation

Thermal control policies

Stop and go

- Simplest thermal policy
  - Halt or severely slow down CPU if temperature exceeds a threshold
  - restart it when temperature falls below the threshold
- High impact on performance
  - Also leads to discontinuous operation
- Still sometimes used as “last resort” secondary policy to prevent thermal runaway



http://www.eecs.harvard.edu/~dbrooks/hpca2001.pdf

Thermal control policies

Using DVFS for thermal control

- Change the operating frequency and voltage at runitme based on a policy
  - Remember their impact on power: $P_{active} = C\ V_{dd}^2\ f$
- Unlike an on-off actuator, allows to fine tune the operating conditions
  - Apply the minimal slowdown needed to keep temperature under control
- More complex policy, needs to decide which frequency to select

Many policies have been proposed

Thermal control policies

Control-theory based DVFS policies

- Treat the thermal control problem as a closed loop control system
- Many control algorithm have been proposed, e.g: PID
- Control algorithm must be executed at a fast rate (1..100ms)
  - to respond to fast thermal dynamics
  - Introduce an overhead due to the CPU time required to run the policy
- Work best when there are many DVFS operating points
  - Excessive quantization of the control output degrade control quality

https://ieeexplore.ieee.org/document/1635942

Thermal control policies

Optimization-based DVFS policies

- Treat the thermal control problem as an optimization problem
- Different objective functions have been proposed
  - Maximize performance subject to temperature constraints
  - Minimize temperature subject to task execution deadline constraints
- Solving optimization problems is CPU-intensive
  - High policy ovehead, hard to operate at the 1..100ms timescale required
  - Sometimes the result can be precomputed and used as a look-up table
- Works best when there are few DVFS operating points to choose from
  - Too many operating points cause combinatorial explosion

Model-predicitive control

- A blend of optimization and control theory

https://ieeexplore.ieee.org/document/4484671

Thermal control policies

Task migration based thermal control policies

- Exploit mult-core architectures
- Main idea: when a core is overheating, move tasks to other cores

Advantages

- Balancing core temperatures does not require to slow down cores

Disadvantages

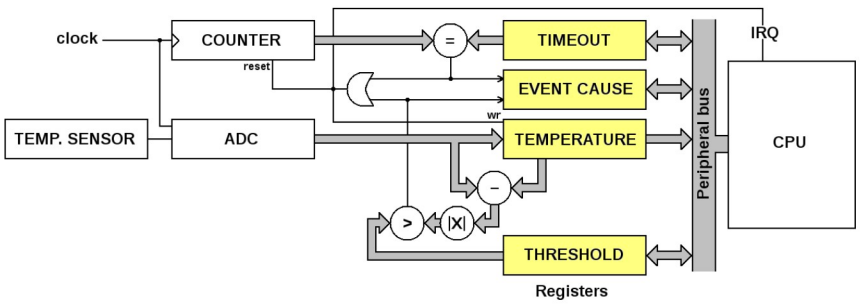
- All cores may be equally hot
  - Secondary policy needed
- Task migration incurs an overhead
  - OS overhead, additional cache misses...

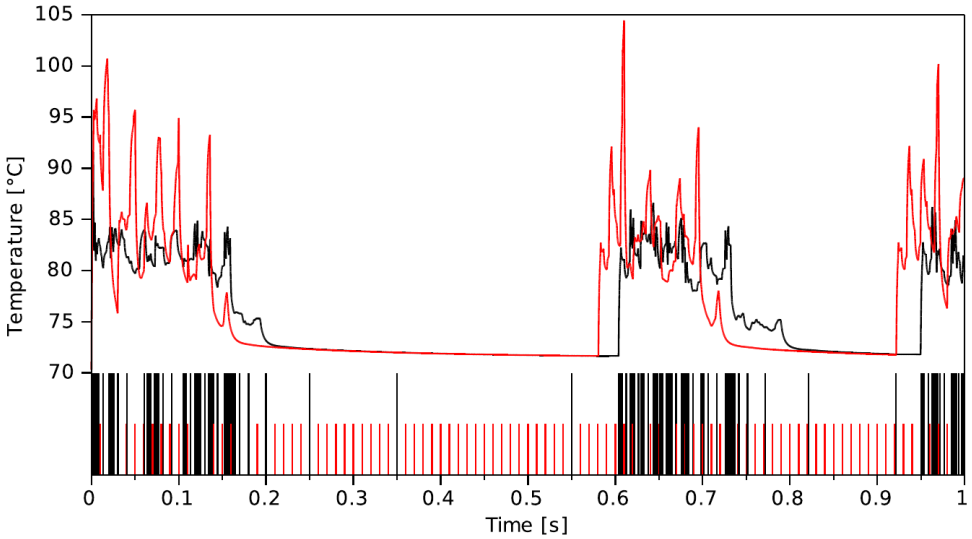
https://engineering.purdue.edu/~vijay/papers/2004/heat-and-run.pdf

Thermal control policies

Event-based control theory

- The policy is not executed at a fixed rate, but when events occur
- Based on a hardware-software codesign
  - A small hardware state machine monitors sensors and generates events
  - Events cause the execution of a software control policy based on control theory
- Faster response with less overhead





https://ieeexplore.ieee.org/document/7890422