Design for Low-Power IoT Systems: Markov Decision Processes and Power/Performance/Thermal (PPT)

Marilyn Wolf, Georgia Tech

Outline

- Non-functional design parameters and system-level methodologies
  - Modeling complexity
  - System design space
- Modeling techniques:
  - WCET
  - Thermal
  - Reliability
- Statistical models for non-functional parameters:
  - Markov decision processes
  - Solution algorithms

Complex application requirements

- Real-time (soft and hard) performance.
  - Throughput and latency.
- Power consumption.
- Thermal performance.
- Complex functionality.

Causes of modeling complexity

- Complex architecture, microarchitecture, logic design.
- Complex physics and multi-physics.
- Process variation.
- Aging-induced variation.

Software performance analysis

- Worst-case execution time (WCET):
  - Worst-case under any inputs or system state.
- Sometimes interested in best-case execution time.
- Software performance analysis:
  - Execution time = f(program path, path timing).
- Path timing is hard:
  - Pipelining.
  - Caches for single process.
  - Multi-tasking cache behavior.

Path analysis

- Exponential number of paths in a general program.
- Identification of arbitrary paths is a halting problem.
- Some methodologies eliminate certain program constructs.

Path timing

- Execution time of an instruction in a modern CPU depends on its context:
  - Other instructions in the pipeline.
  - Cache.
- Pipeline issues:
  - Pipeline stalls.
  - Superscalar issue.
- Bounds on paths can be found implicitly.

Cache behavior

- Memory access time depends on cache state.

- Cache state depends on entire history of program.

tag	contents
tag	contents
tag	contents
tag	contents

Cache and multitasking

- Each task has its own cache behavior.

- Tasks interact in the cache.

task 2

conflict

task 1

Power and energy

- Dynamic power consumption:
  - Energy consumed while performing useful work.
  - $E = \frac{1}{2}CV^2$
  - Ideal gate consumes zero power while idle.

- Static power consumption from multiple mechanisms:
  - Short-circuit current.
  - Leakage.
  - Consumes power even while idle.

- How to reduce power:
  - Reduce power supply voltage to lower dynamic power.
  - Remove power supply to remove dynamic and static power.

Dynamic voltage and frequency scaling

- Power consumption $P \propto V^2$.
- Delay $\delta \propto \frac{1}{V}$.
- Reducing power supply voltage decreases power at a faster rate than it increases delay.

- DVFS monitors activity, selects power supply voltage and clock frequency based on performance demands.

Power management algorithms

- Dynamic power only---slow down the processor to just make the deadlines.

- Only two power supply voltages are required to optimize power consumption.

- Race-to-dark runs as fast as possible, then shuts down to minimize leakage.

Computer system energy consumption

Component	Power (W)
CPU	100-200
Memory	25
Disk	10-15
Board	40-50
Power/fans	30-40
Total	200-350

server
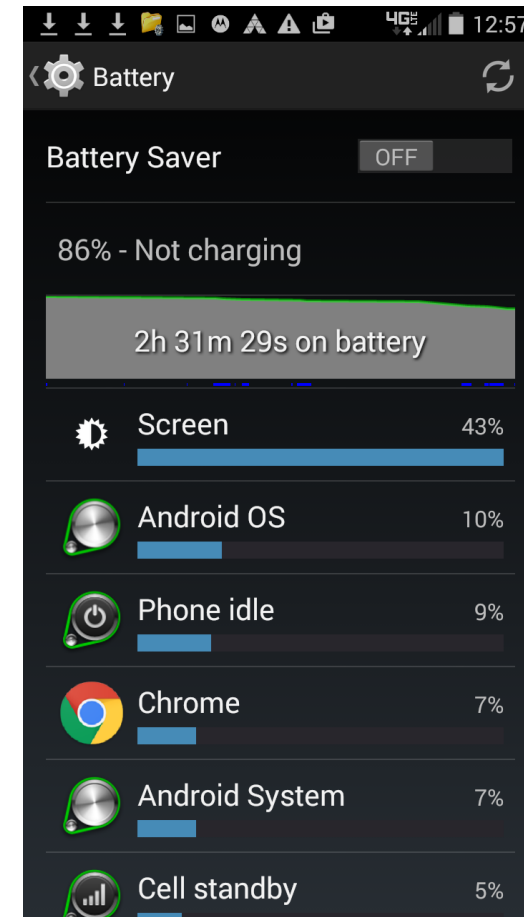
smartphone

Current requirements

- Intel Xeon E7-8800:
  - $I_{CC\_MAX} = 120\ A$
  - Operating voltage of 1.3V

- Craftsman Arc Welder:
  - $I = 60\ A$
  - Operating voltage of 120V.

Welder draws higher power but CPU has impressively high current density.

Heat transfer mechanisms

- Conduction.
  - Molecular motion.
- Radiation.
  - Electromagnetic energy.
- Convection.
  - Bulk fluid motion.

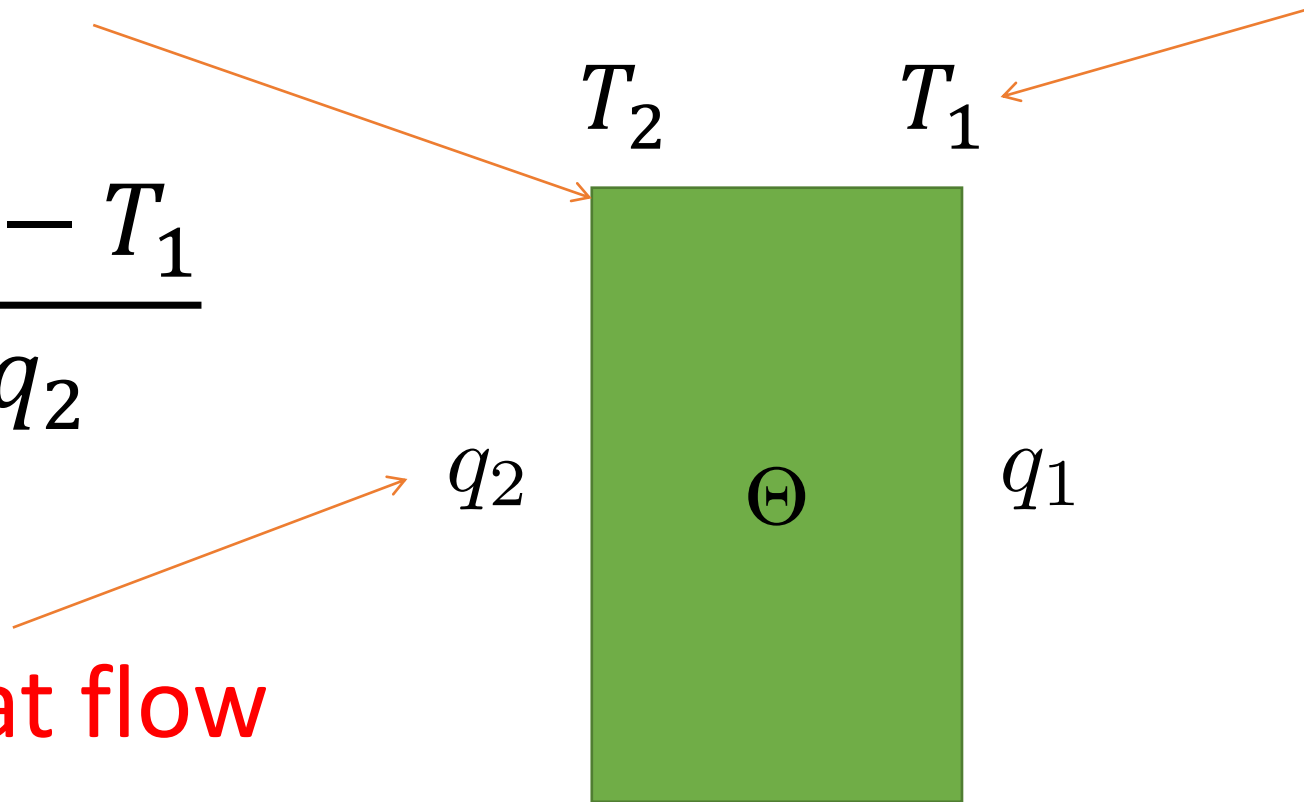Heat carried through a solid.

Can be transmitted in a vacuum.

Air or water flow.

Key thermal ratings

- Transistor junctions must be kept below maximum junction temperature:
  - $T_{J,max} = 85°C$
  - Temperature at which heat damages the transistor structures.
- Chip specifies thermal design power (TDP).
  - Amount of operating heat that its cooling system must be able to dissipate.

Thermal drives performance

TDP ➡️ P ➡️ f
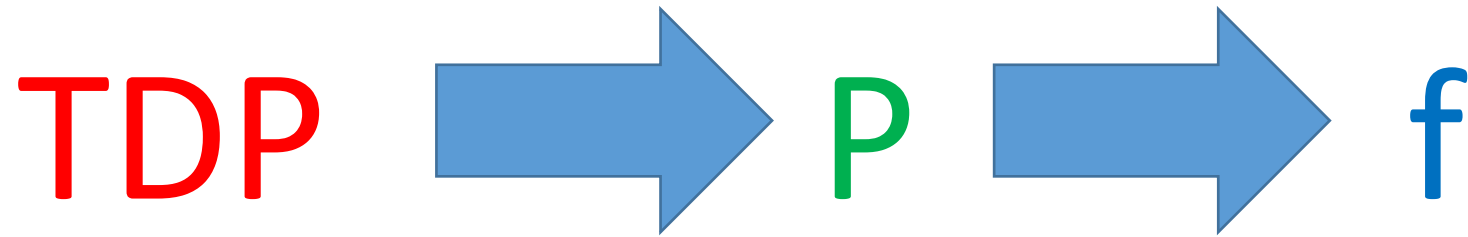
Physical properties

- Specific heat:
  - Relationship between heat input/output and temperature.
  - Measured in Joules/kilogram-Kelvin.
- Thermal conductivity:
  - Relationship between temperature difference and heat flow per unit time.
  - Measured in Watts/meter-Kelvin.

material	specific heat (J/kg K)	thermal conductivity (W/m K)	density (kg/m³)
silicon	710	149	2.3×10^{-3}
ceramic (aluminum nitride)	740	150	3.3×10^{-3}
carbon steel	620	41	7.9×10^{-3}

Thermal properties of objects

- Thermal resistance R:
  - Thermal conductivity for a specific shape and size of material.
  - $R = \frac{l}{kA}$, length l, area A, thermal conductivity k.

- Thermal capacitance C:
  - Specific heat for a specific shape and size of material.
  - $C = mC_m$, mass m, specific heat C_m.

Thermal/electrical analogy

electrical	thermal
charge Q	thermal energy Q
current I	heat flow P
voltage V	temperature T
resistance R	thermal resistance R
capacitance C	thermal capacitance C

Physical laws of thermal behavior
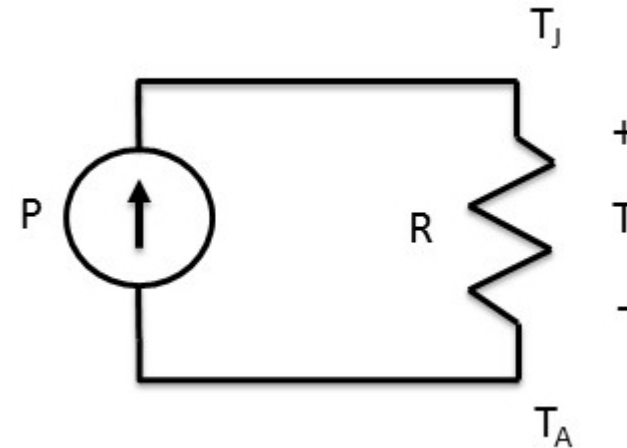
- Fourier's Law of Heat Conduction:
  - $T = PR$
- Newton's Law of Cooling:
  - $P = C\dfrac{dT}{dt}$

Steady-state temperature

- Use thermal resistance to determine steady-state temperature.
  - Calculate temperature difference from ambient to junctions.
- Thermal circuit has a heat source P, thermal resistance R.
- Output temperature is measured across thermal resistance.
  - $T_J = T_A + P\mathrm{R}$
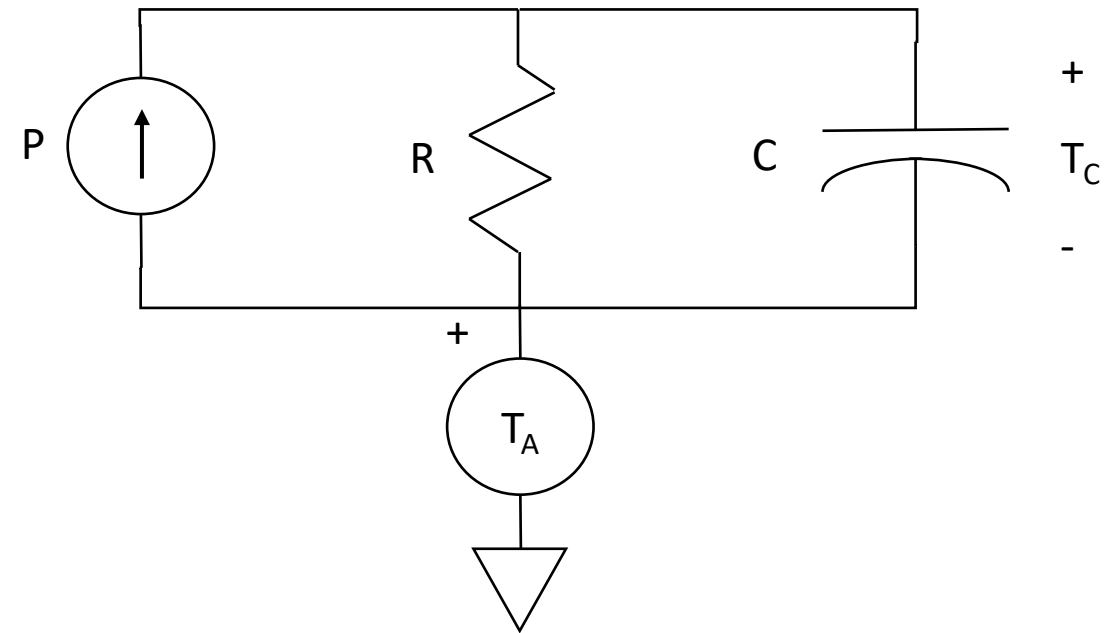
Example: heat sink performance

- Computer power $P = 20\ W$.
- Ambient temperature 20℃.

- Case 1---no heat sink:
  - $T_{none} = 20 + 20W \cdot 10\frac{℃}{W} = 220℃$

- Case 2---heat sink:
  - $T_{sink} = 20 + 20W \cdot 1.5\frac{℃}{W} = 50℃$

Thermal transient analysis

- Chip and heat sink form a thermal RC circuit.
  - Measure chip temperature relative to ambient.
- Temperature as a function of time:
  - $T(t) = (T_0 - PR)e^{-t/RC} + PR + T_A$
  - Temperature above ambient at $t = \infty$ is PR.
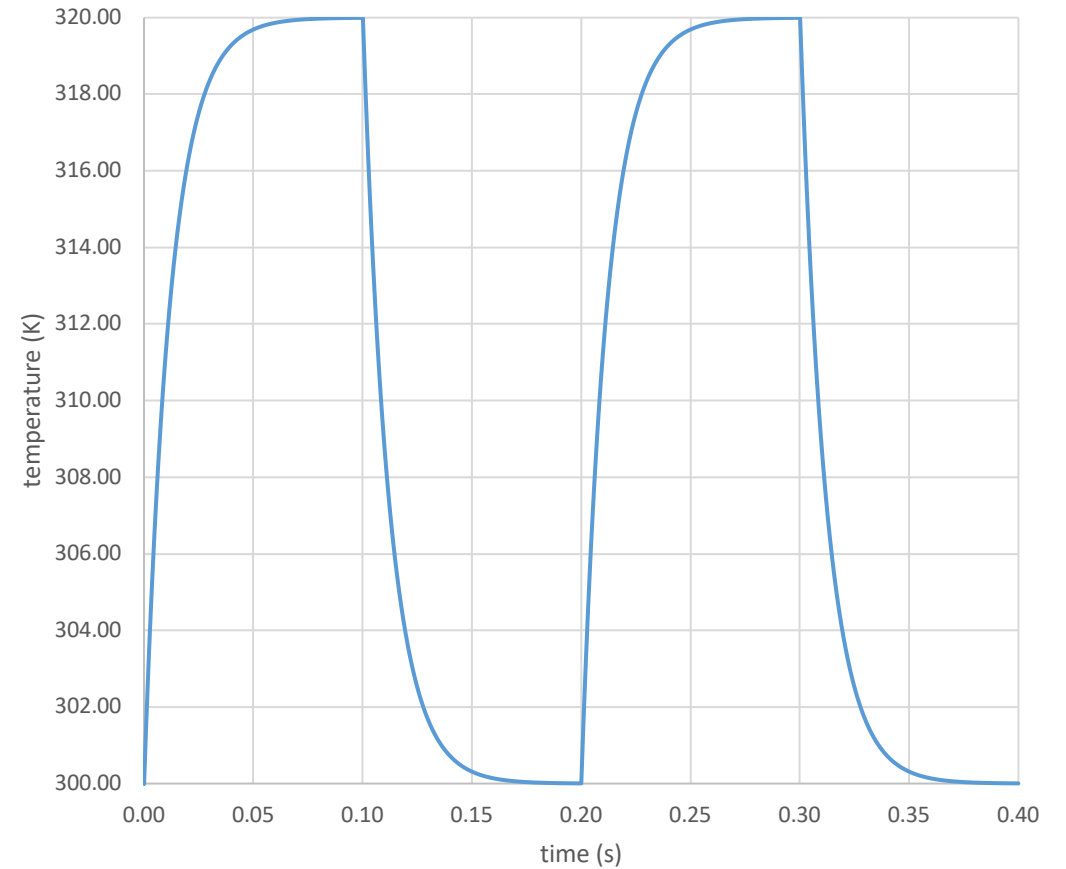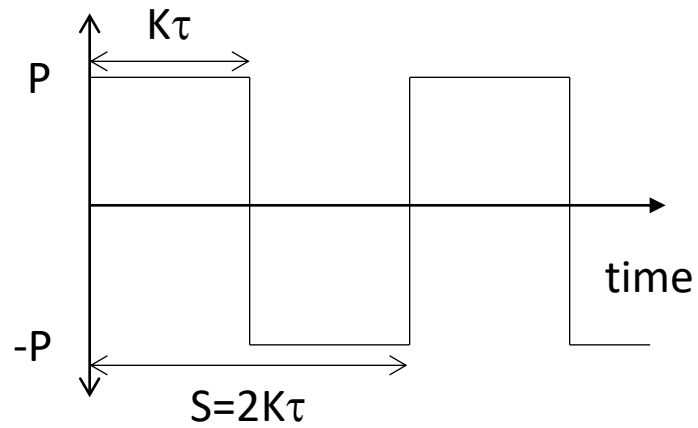
Example: thermal RC model of chip temperature

parameter	value
R	0.5 K/W
C	0.03 J/K
P	50 W
T_0	0 K
T_A	300 K

- $T(t) = 325 - 25e^{-\frac{t}{0.015}}$
- Thermal time constant $0.015\ sec$.
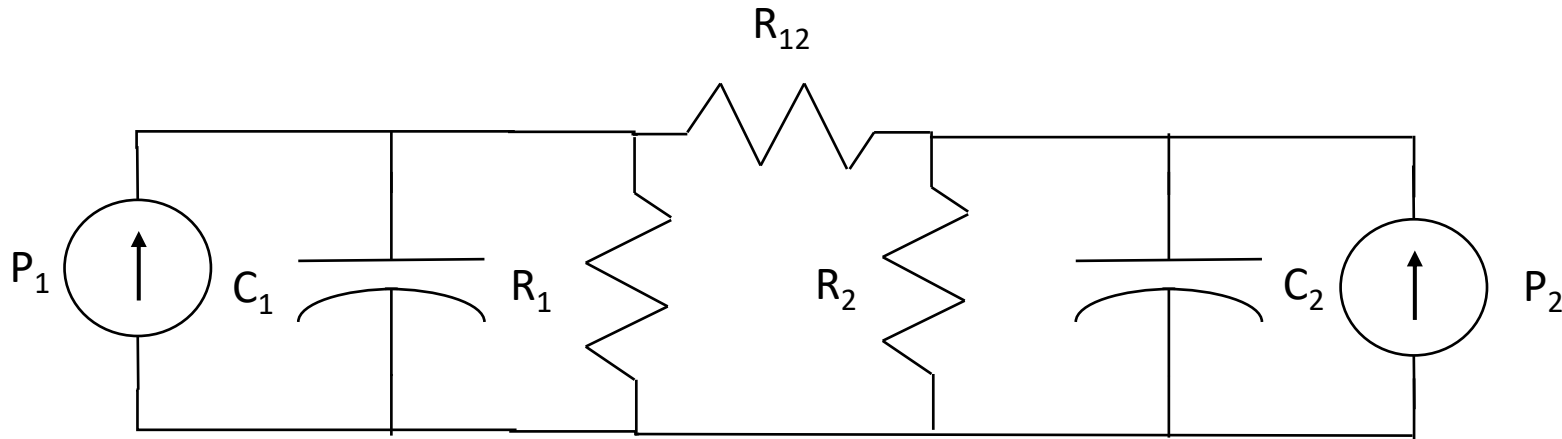- Steady-state temperature $325\ K$.
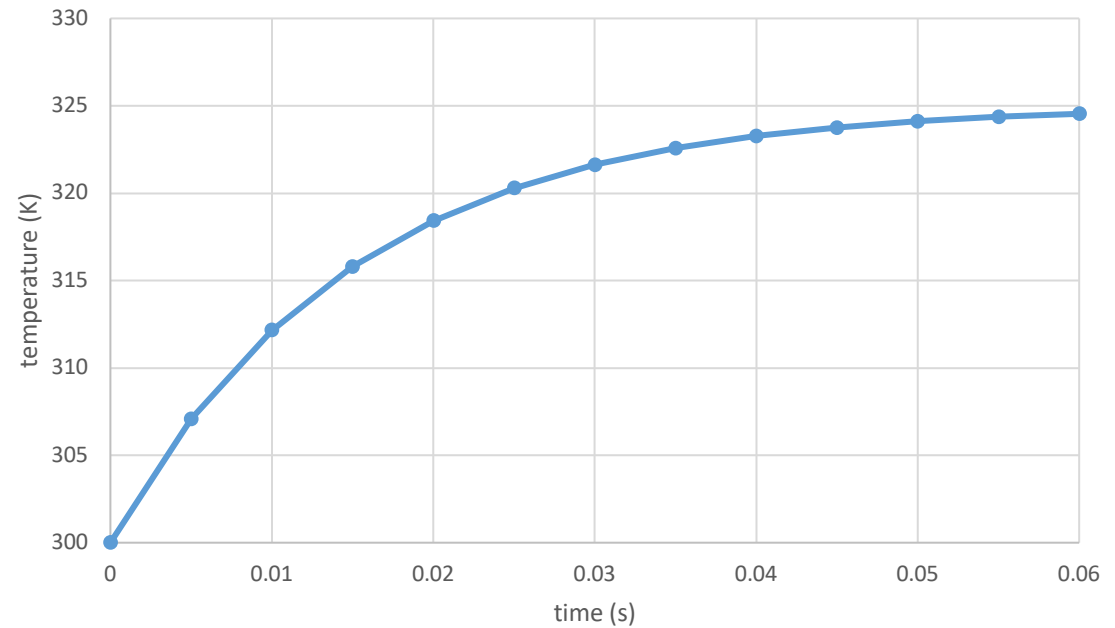
Thermal square wave

- Chip runs periodically:

Dual-core processor

- Two processors alternate between run, stop.
- Cores are connected by thermal resistance.

Dual-core thermal analysis

- Can borrow a result from electrical circuits:
  - Assume 50% duty cycle, period $S = 2K\tau$.

- Upward and downward temperature waveforms:
  - $T^u(t + t_0) = \left(-T_p - P\right)e^{-t/RC} + P$
  - $T^d(t + t_0) = \left(T_p + P\right)e^{-t/RC} - P$

- Temperature cycles between $T_p, -T_p$:
  - $T_p = \left(-T_p - P\right)e^{-K} + P = \left(T_p + P\right)e^{-K} - P$

- So $\dfrac{T_p}{P} = \dfrac{1 - e^{-K}}{1 + e^{-K}}$

Dual-core thermal behavior
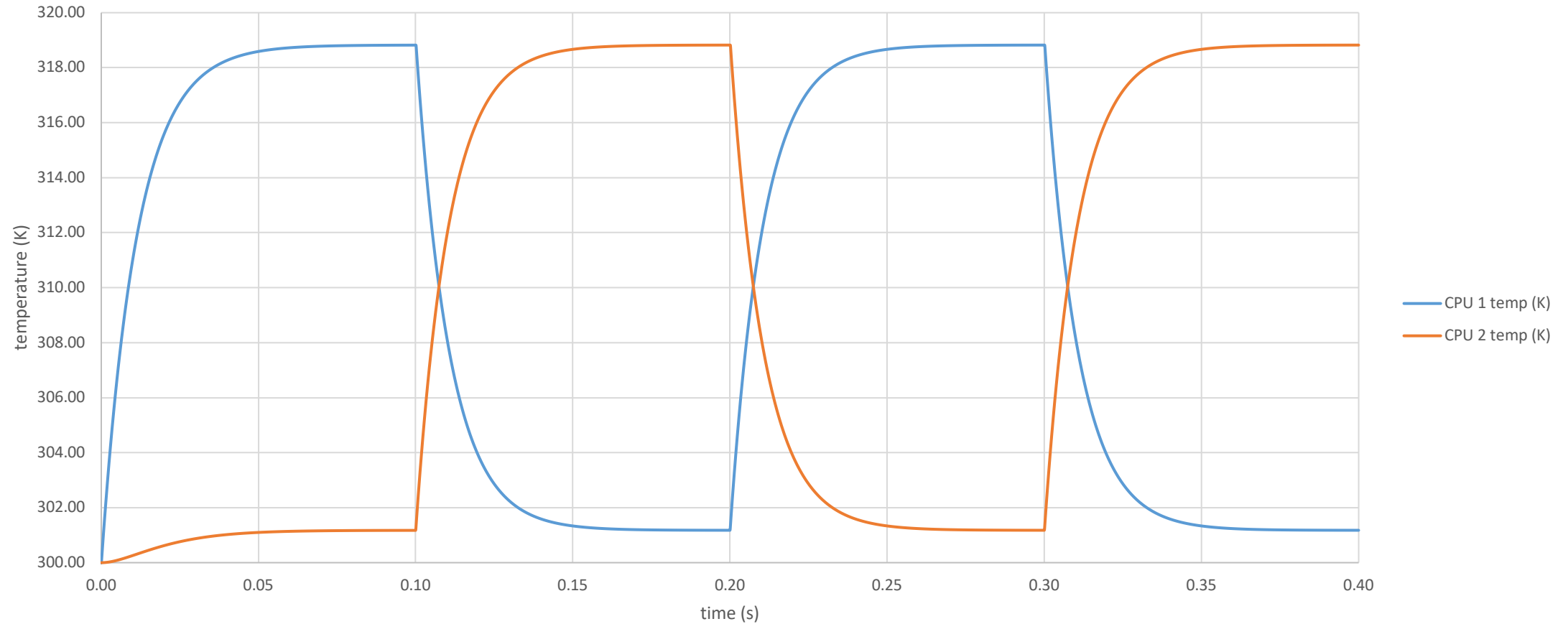
Heat and reliability

- Heat contributes to aging.
  - Higher temperatures cause chips to fail earlier.
- Arrhenius' equation describes the relationship between energy and the rate of physical processes:
  - $r = Ae^{-E_a/kT}$
  - Activation energy E_a is determined by energy required to promote electrons to high orbits.
  - Arrhenius prefactor A is measured experimentally.

Electromigration

- Electromigration is a common temperature-related failure mechanism.
    - Heat causes some molecules in wire to release free atoms.
    - Current flowing through wire causes free atoms to move.
    - Destructive feedback---thinner wire segments heat more, causing more rapid failure.
- Failure rate modeled using Black's equation:
   - $MTTF = AJ^{-n}e^{E_a/kT}, 1 \leq n \leq 3.$

Lifetime analysis

- Chip temperature often varies over time based on use case and computing activity.

- We can model aging as a function of temperature:
  - $R(t) = \frac{1}{kT(t)} e^{-E_a/kT(t)}$
  - $\varphi_{th} = \int_0^t \frac{1}{kT(t)} e^{-E_a/kT(t)}$

- Chip-level reliability engineering:
  - Minimize hot spots.
  - Use operating system to spread workload across cores.

Thermal management

- A combination of hardware and software is used to manage thermal behavior.

- On-chip temperature measured using band gap reference circuit.

- Processor may provide a software interface to on-chip temperature sensors.
  - Intel Thermal Monitor 1 turns the clocks off and on at a duty cycle chosen for the processor type, typically 30%-50%.
  - Intel Thermal Monitor 2 uses dynamic voltage and frequency scaling mechanisms to reduce both the clock speed and power supply voltage of the processor.

Markov decision processes
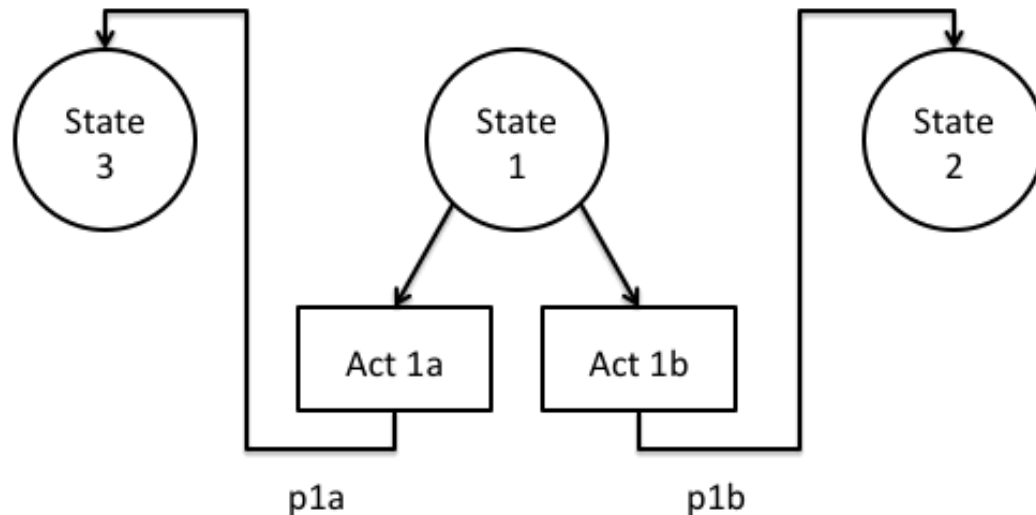
- Probabilistic transitions combined with inputs.
  - Given an input at a state, next state is chosen probabilistically.
- A policy π defines the actions in each state s.
  - Optimal policy maximizes rewards.

© 2018 Marilyn Wolf

MDP model

- States S.

- Actions S.

- Probability that action a in state s gives transition to s' $P_a(s, s')$.

- Reward for action $R_a(s, s')$.

- Discount factor γ.

- Find policy π to maximize time-discounted reward:
  - $\sum_{t \geq 0} \gamma^t R(s_t, s_{t+1})$

Value iteration

- $V_{i+1}(s) = \max_a [\sum_{s'} P_a(s, s')\{R_a(s, s') + \gamma V_i(s, s')\}]$
- Value at each step is maximum over all possible actions.
- Iterate until converged.

Policy iteration

1. Find policy $\pi(s) = \arg max_a \, P_a(s, s')\{R_a(s, s') + \gamma V_i(s, s')\}$

2. Iterate until converged $V \ (s) = \sum_{s'} P_{\pi(s)} (s, s')\{R_{\pi(s)}(s, s') + \gamma V_i(s, s')\}$

3. Repeat 1-2 until converged.

Reinforcement learning

- Identify transition probabilities using random search.
  - Explore new space while making use of learned model.
- Step at time t:
  - Agent is in state s_t.
  - Observe environment o_t, reward r_t.
  - Choose action a_t.
  - Transition to state s_{t+1}.