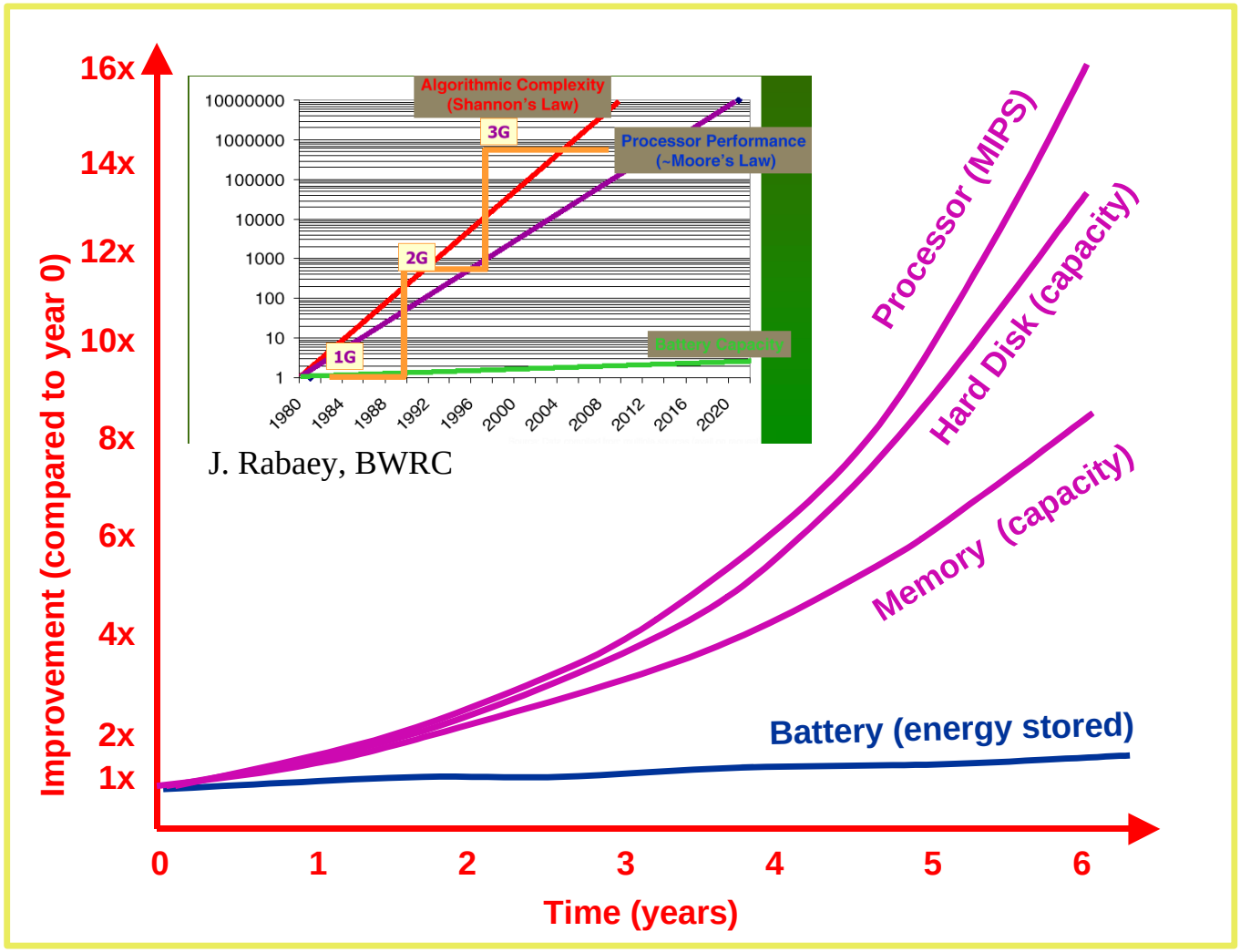
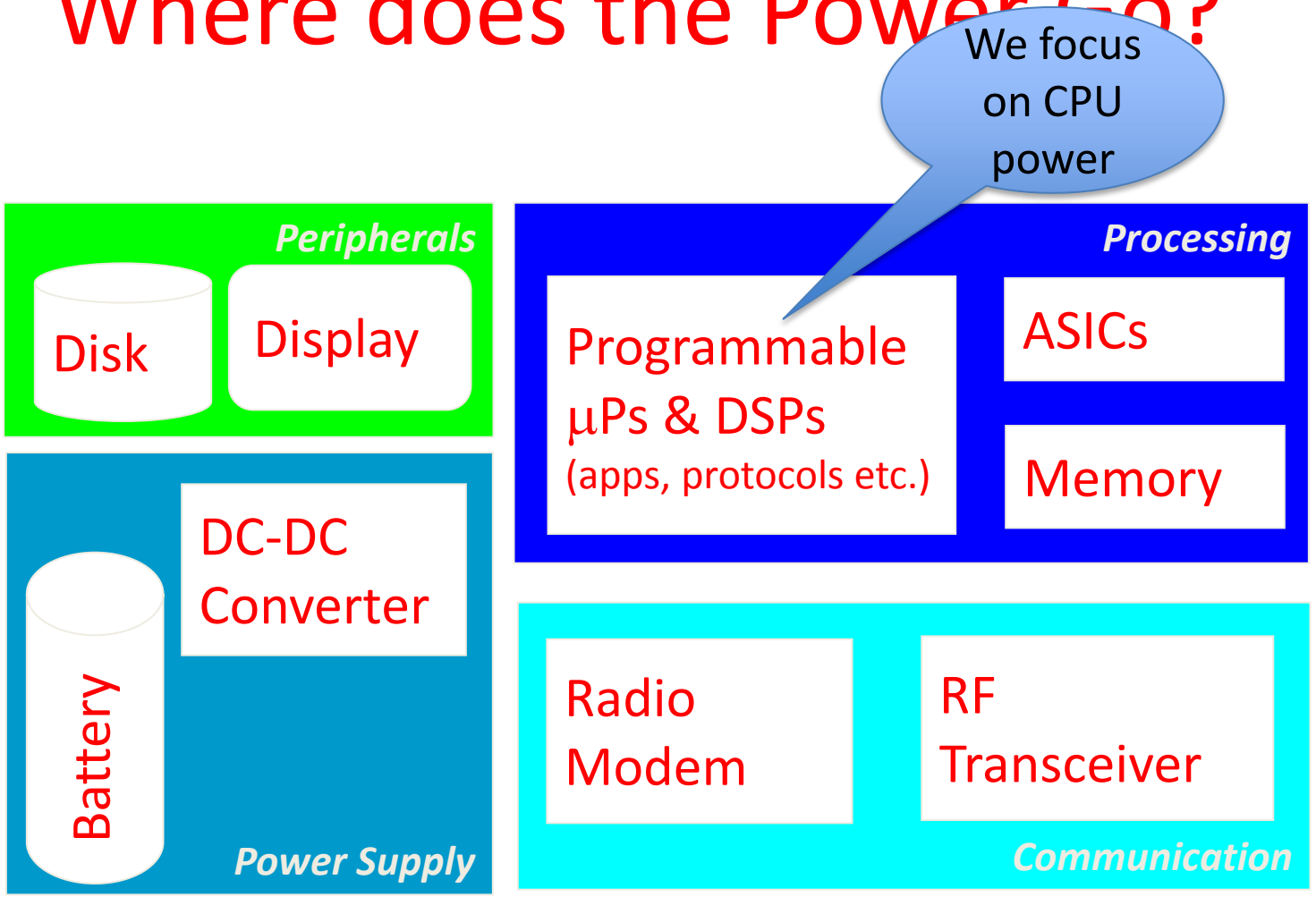
Power Management with DVFS

Instructor:
Zonghua Gu

The Case for Power Awareness



Where does the Power Go?



Leakage vs. Dynamic

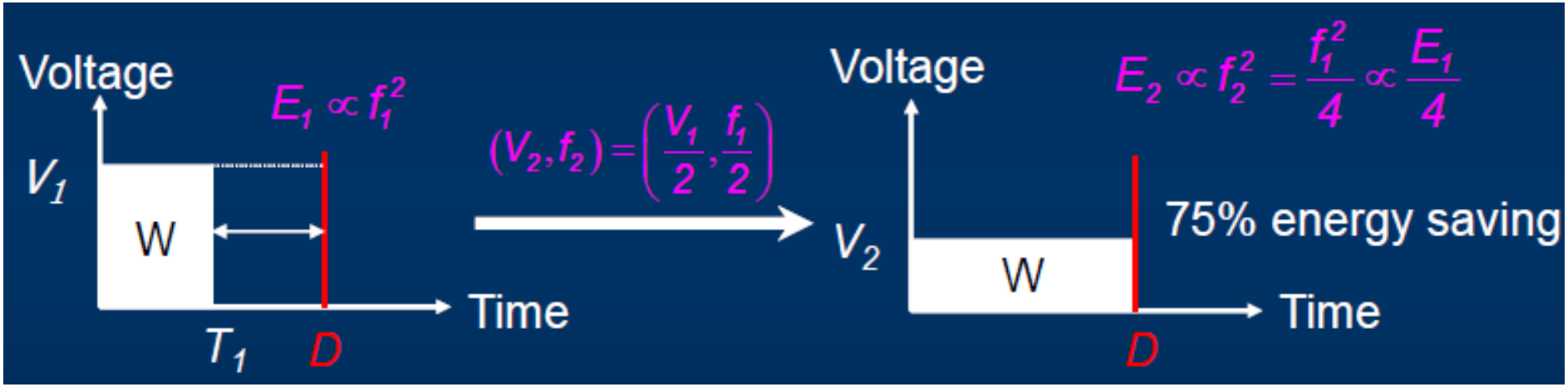
- Leakage power (static power) is consumed even if the circuit is idle (standby mode). The only way to avoid it is to turn off the CPU.
  - The only way to reduce leakage power is to turn off the power supply
- Dynamic power is still the main source of power consumption.
  - In the past, Leakage power was considered negligible compared to dynamic power.
  - Today, Leakage power is approaching or exceeding size of dynamic power.

Dynamic Power/Energy

- Power consumption： $P=\alpha*C*V^2*f$
  - α: switching factor
  - C: effective capacitance
  - V: operating voltage
  - f: operating frequency
- Energy consumption for task with execution time t_{exe} (it is obvious that $t_{exe} \propto 1/f$)：
  - $E=P*t_{exe}=\alpha*C*V^2*f*t_{exe}=\alpha*C*V^2*N_{cycles}$
  - N_{cycles}: total execution cycles
- Assuming $f \propto V$ (approximation):
  - $P \propto V^3 \propto f^3$; $E \propto V^2 \propto f^2$

Reducing Power Consumption

- Given power consumption P=α*C*V^2*f, we can reduce P by:
  - At device level:
    - reduce capacitance C
  - At system level:
    - Reduce CPU voltage V
    - Reduce CPU frequency f



Power vs. Energy

- Which is more important?
- Sometimes we care about power consumption:
  - heat dissipation, cooling
  - physical deterioration due to temperature.
- Sometimes we care about total energy consumption:
  - battery life for portable devices
  - power bills (money) for large servers

Dynamic Voltage and Frequency Scaling (DVFS)

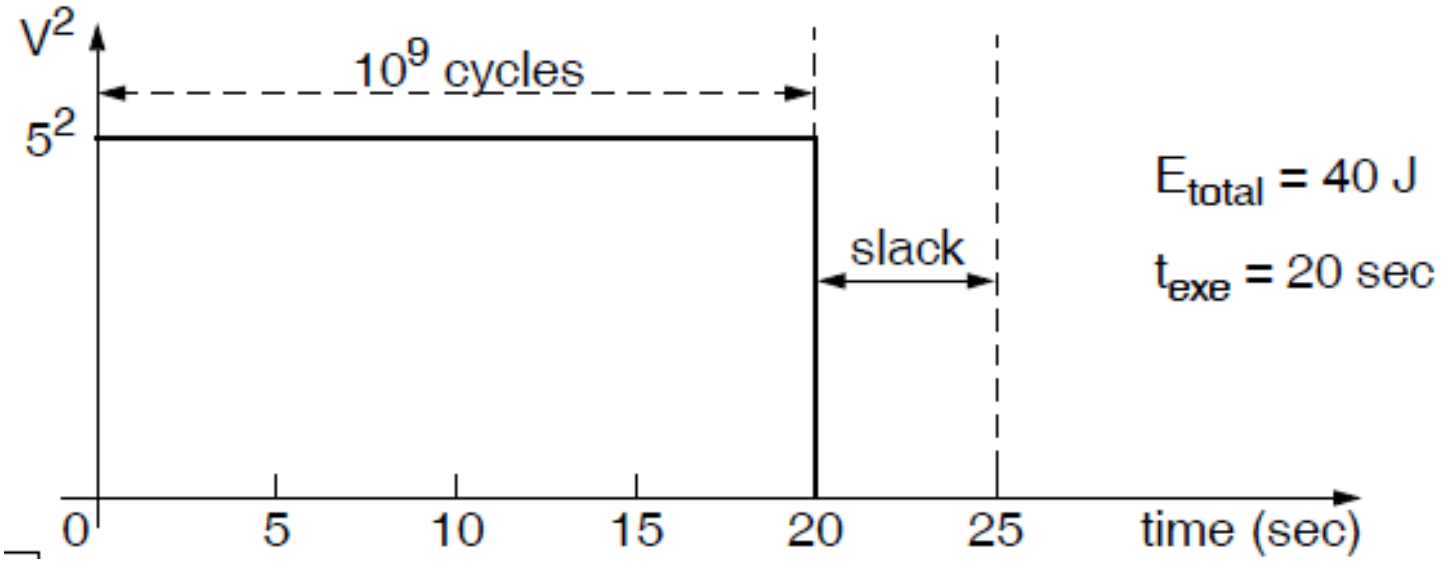
- A device can be run at different speeds at several different (supply voltage, CPU frequency) pairs
- Supply voltage can be chosen by the OS through execution of particular instructions. CPU frequency is automatically adjusted to fit the current supply voltage.
- Execution of jobs can be slowed down to save power as long as all jobs are completed by their deadline.

Scheduling w/ DVFS

- Reducing supply voltage increases execution time!
- The scheduling problem:
  - Which task to execute at a certain moment on a certain processor, *and at which voltage level, so that time constraints are fulfilled and energy consumption is minimized?*

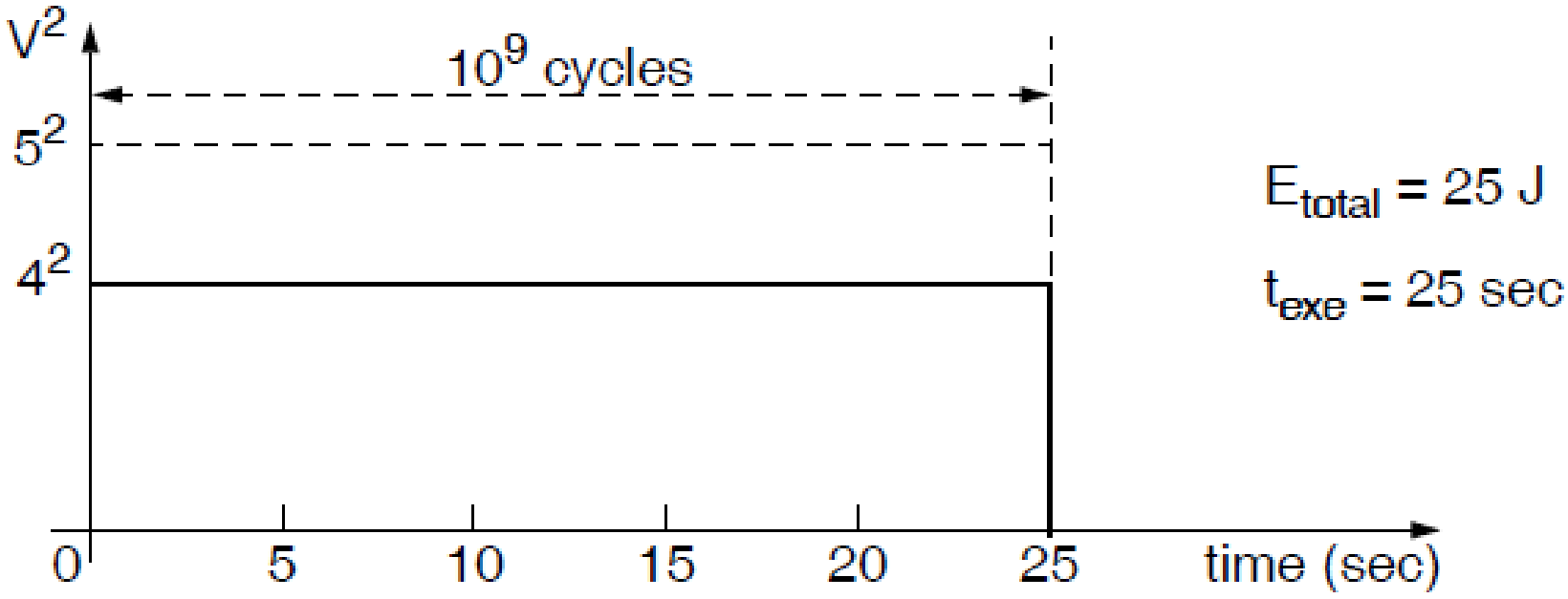
Example

- Consider a task τ, and CPU runs at full speed until τ finishes:
  - Processor nominal (maximum) voltage: 5V.
  - Total computation: 10^9 execution cycles; deadline: 25 seconds.
  - CPU frequency: 50MHz (50 million cycles/sec) at nominal voltage; Energy per cycle: 40 nJ/cycle at nominal voltage.
  - Execution time t_{exe}= 10^9 (cycles)/50*10^6 (cycles/sec)=20 s
  - Total energy E_{total} = 10^9 (cycles) * 40 (nJ/cycle)=40 J



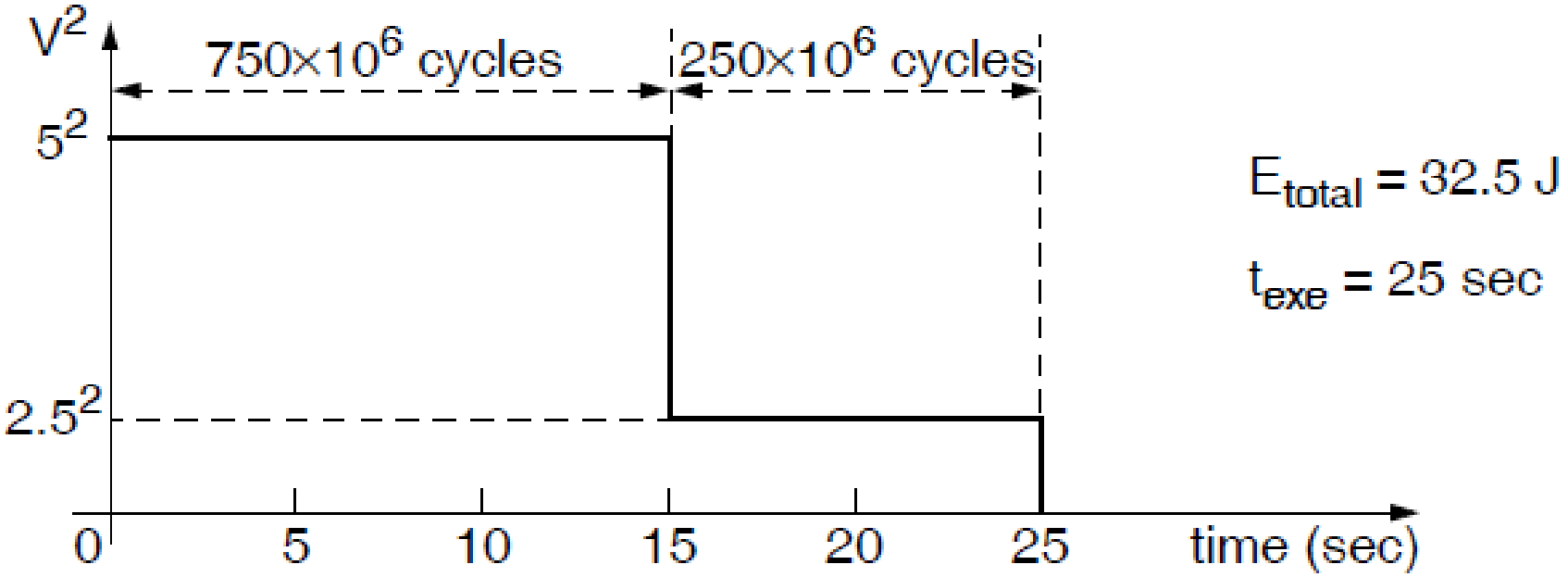
Reduced Voltage & Frequency

- Reduce voltage at time 0 to V = 4V
  - CPU frequency: 50*(4/5) = 40MHz； Energy per cycle: $40*(4/5)^2$ = 25nJ/cycle.
  - Execution time t_{exe}= 10^9(cycles)/40*10^6 (cycles/sec)=25 s
  - Total energy E_{total}= 10^9 (cycles) * 25 (nJ/cycle) = 25 J



Reduced Voltage & Frequency

- Reduce voltage at time 15 to V = 2.5V
  - CPU frequency: 50*(2.5/5) = 25MHz； Energy per cycle: $40*(2.5/5)^2$ = 10nJ/cycle.
  - Execution time t_{exe}=$750*10^6$(cycles)/$50*10^6$ (cycles/sec) + $250*10^6$(cycles)/$25*10^6$ (cycles/sec) =15+10=25 s
  - Total energy E_{total}=$750*10^6$ (cycles) * 40 (nJ/cycle) + $250*10^6$ (cycles) * 10 (nJ/cycle)=32.5 J



Basic Principle

- If a processor uses a single supply voltage and completes a program at exactly the deadline moment, the energy consumption is minimized.

[Zhang10]

- Xiao Zhang et al, An Evaluation of Per-Chip Non-Uniform Frequency Scaling on Multicores, 2010
- *Similarity grouping*:
  - On a multi-chip system, group applications with similar last level cache (LLC) miss ratios to run on the same multicore processor chip, in order to lower the voltage/frequency of the chip running memory-intensive tasks w/ high LLC miss ratios.
- Question: this seems opposite to [Blagodurov11] paper, which tries to distribute memory-intensive tasks on different chips?
  - [Blagodurov11] addresses NUMA, while [Zhang10] addresses UMA. But even for UMA, since each chip has its private LLC cache, the cache contention issue is similar.
  - [Zhang10]'s approach exacerbates cache contention on the chip running mem-intensive tasks, but power consumption is reduced by reducing voltage/frequency of the the chip running mem-intensive tasks.
  - [Blagodurov11] focuses on performance and does not consider power consumption.

[Zhang10]

- Question from audience
  - On a multi-chip system, tasks running on different chips may interfere with each other. Even though they do not share CPU, the share the bus, interconnect, memory controller, etc.
- The Frequency-to-Performance Model did not consider this interference, hence may not be very accurate.