Energy-Aware Computing Systems

Energiebewusste Rechensysteme

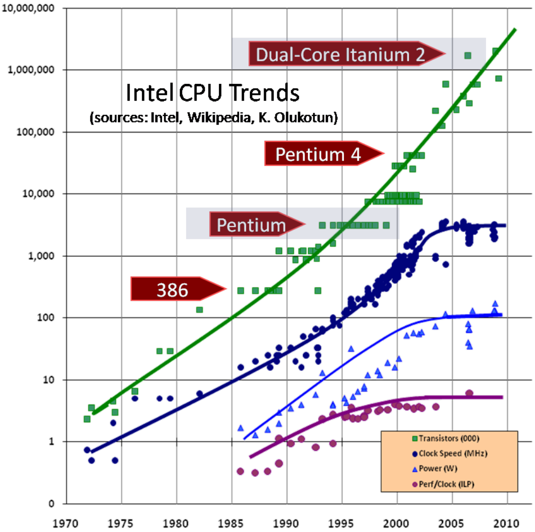
II. Principles

Timo Hönig

2019-05-02

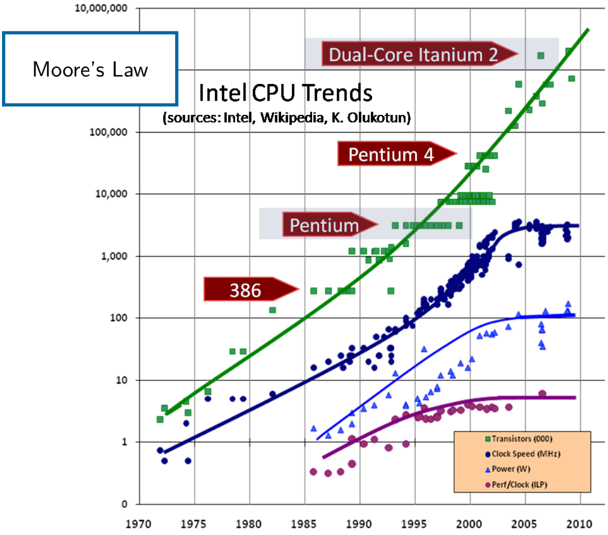
Agenda

Preface: The Free Lunch is Over



Sutter '05 [7]

Preface: The Free Lunch is Over

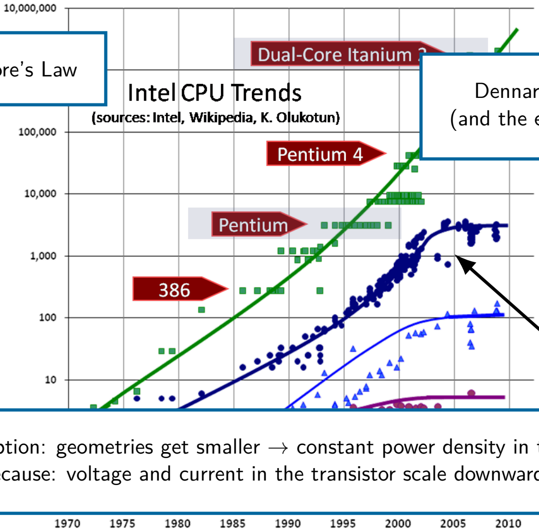


Sutter '05 [7]

Preface: The Free Lunch is Over

Moore's Law

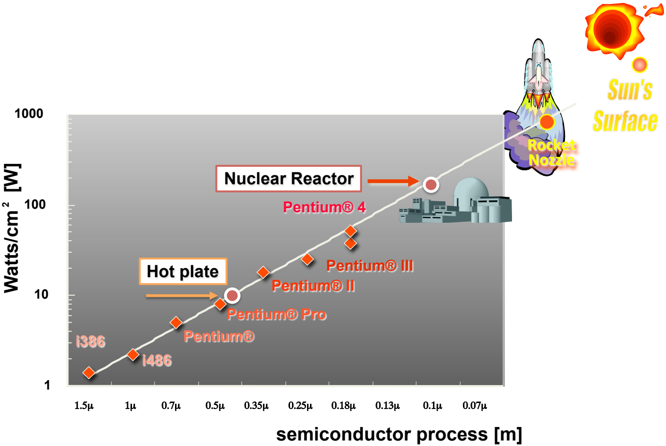
Dennard Scaling
(and the end thereof)



assumption: geometries get smaller → constant power density in transistors
because: voltage and current in the transistor scale downward, too

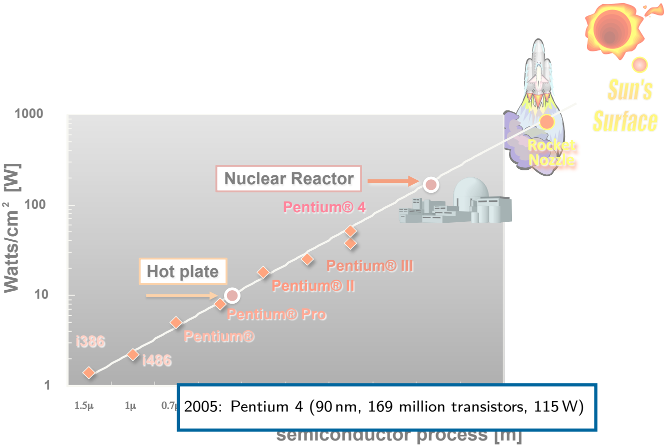
Sutter '05 [7]

Preface: The Power Wall



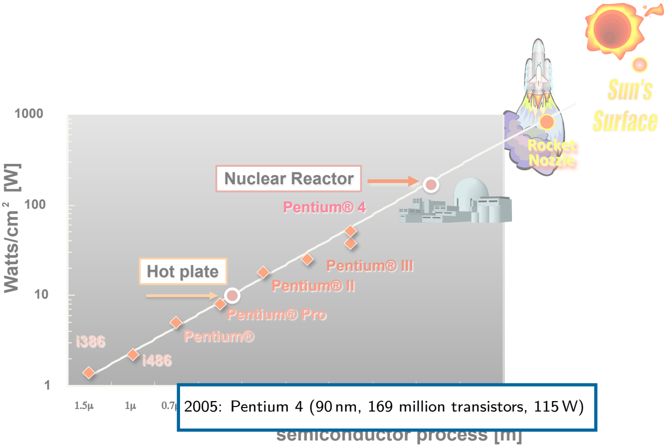
Pollack '99 [5]

Preface: The Power Wall



2005: Pentium 4 (90 nm, 169 million transistors, 115 W)

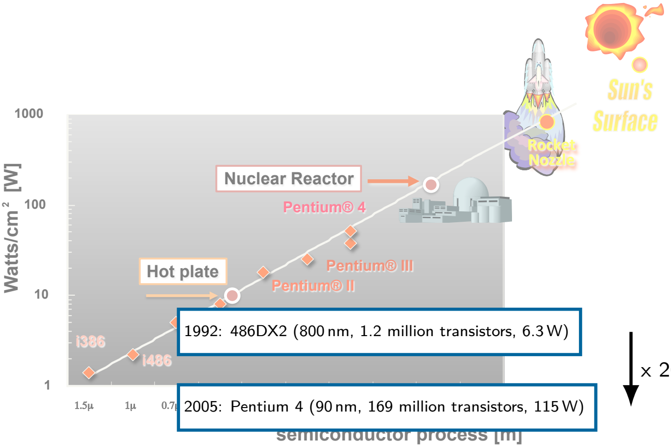
Preface: The Power Wall



2005: Pentium 4 (90 nm, 169 million transistors, 115 W)

2019: Core i9-7900X (14 nm, approx. 7 **billion** transistors, 140 W)

Preface: The Power Wall



1992: 486DX2 (800 nm, 1.2 million transistors, 6.3 W)

x 20

2005: Pentium 4 (90 nm, 169 million transistors, 115 W)

x 1.2

2019: Core i9-7900X (14 nm, approx. 7 **billion** transistors, 140 W)

Agenda

Disambiguation: Energy-Aware Computing Systems

recap: meaning of the lecture labelling in linguistic terms:

en·er·gy (gr.) *energeia*: word based upon *ergon*, meaning *work*

1. capacity for the exertion of power
2. a fundamental entity of nature that is transferred between parts of a system in the production of physical change within the system

aware (old en.) *gewær*

1. having or showing realization, perception, or knowledge
2. state of being conscious of something

com·put·ing (lat.) *computare*: *com* (together) + *putare* (to settle)

1. task of making a calculation
2. to use a computer

sys·tems plural of (gr.) *systēmas*: to place together

1. a regularly interacting or interdependent group of items forming a unified whole
2. a group of devices (...) or an organization forming a network especially for distributing something or serving a common purpose

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems
energy	**efficient**	computing	systems

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems
energy	**efficient**	computing	systems
power	aware	computing	systems

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems
energy	**efficient**	computing	systems
power	aware	computing	systems
power	**efficient**	computing	systems

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems
energy	**efficient**	computing	systems
power	aware	computing	systems
power	**efficient**	computing	systems

- energy vs. power

energy : capacity to do work
power : rate of doing work

- to be aware as a prerequisite to be efficient

aware : perception and sensing → e.g., measure ground truth
efficient : retrospective, current, and predictive → e.g., ↑ results, ↓ efforts

Disambiguation: Energy-Aware Computing Systems

- dissecting the terminology

energy	aware	computing	systems
energy	**efficient**	computing	systems
power	aware	computing	systems
power	**efficient**	computing	systems

- energy vs. power

energy : capacity to do work
power : rate of doing work

- to be aware as a prerequisite to be efficient

aware : perception and sensing → e.g., measure ground truth
efficient : retrospective, current, and predictive → e.g., ↑ results, ↓ efforts

- also consider and reflect on: efficient vs. effective

efficient : useful work per quantity of energy invested
effective : degree of reaching a pursued goal

Energy-Aware Computing Systems

- leading questions → system constraints
  - what is the average or maximum power demand? → supply requirements
  - which limits (e.g., thermal) must be adhered to? → demand limit
  - is there a maximum energy demand? → extend system service duration

Energy-Aware Computing Systems

- leading questions $\rightarrow$ system constraints
  - what is the average or maximum power demand? $\rightarrow$ supply requirements
  - which limits (e.g., thermal) must be adhered to? $\rightarrow$ demand limit
  - is there a maximum energy demand? $\rightarrow$ extend system service duration

- metrics
  - what are the correct **metrics** to answer the leading questions?
  - what correlation towards other (non-functional) system properties must be respected?
  - what are the influencing factors and variables?

Energy-Aware Computing Systems

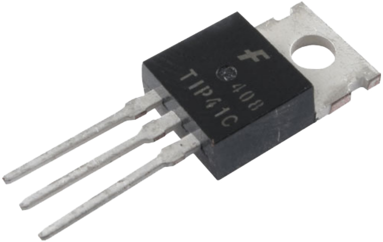
- leading questions → system constraints
  - what is the average or maximum power demand? → supply requirements
  - which limits (e.g., thermal) must be adhered to? → demand limit
  - is there a maximum energy demand? → extend system service duration
- metrics
  - what are the correct **metrics** to answer the leading questions?
  - what correlation towards other (non-functional) system properties must be respected?
  - what are the influencing factors and variables?
- methods
  - what are the correct **methods** to answer the leading questions?
  - how to determine the relevant base data (e.g., power and energy demand)?
  - what is the correct momentum of analysis? → a priori / at runtime / a posteriori

Agenda

Switching Circuit

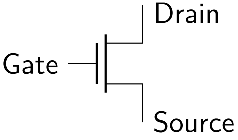
dt. Schaltkreise

- switch: a device for making **and** breaking the connection in an *electric* circuit
- basic components in CMOS technology
  - transistors (*imperfect* switches)
  - wires (interconnect)
- transistor types
  - NMOS (n-type transistor)
  - PMOS (p-type transistor)



Basic System Components: Transistors

- NMOS

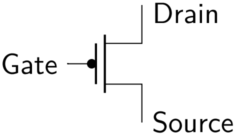


if gate is at logic 1
then the switch is on

- source and drain are connected
- electric current flows

logic 1 $V_{gate} > V_{threshold\_high}$

- PMOS



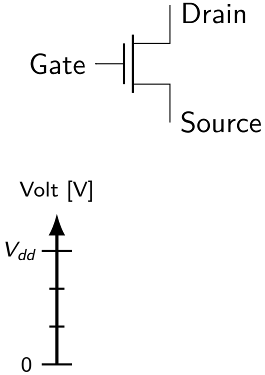
if gate is at logic 0
then the switch is on

- source and drain are connected
- electric current flows

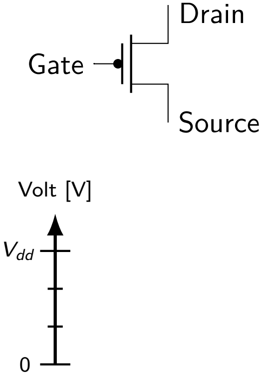
logic 0 $V_{gate} < V_{threshold\_low}$

Basic System Components: Transistors

- NMOS



- PMOS

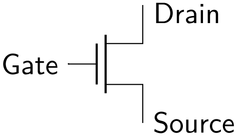


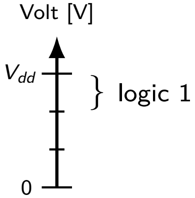
logic 1 $V_{gate} > V_{threshold\_high}$

logic 0 $V_{gate} < V_{threshold\_low}$

Basic System Components: Transistors

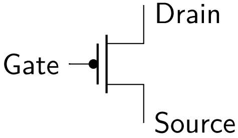
- NMOS

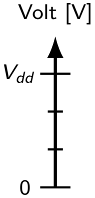




logic 1 $V_{gate} > V_{threshold\_high}$

- PMOS

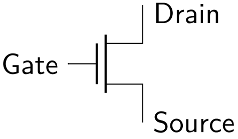


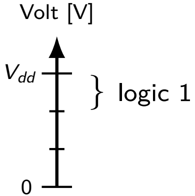


logic 0 $V_{gate} < V_{threshold\_low}$

Basic System Components: Transistors

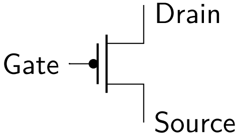
- NMOS

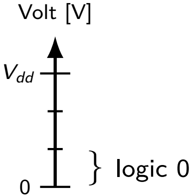




logic 1 $V_{gate} > V_{threshold\_high}$

- PMOS

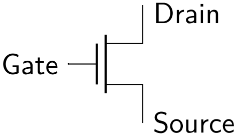




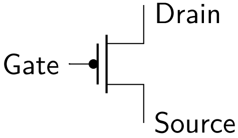
logic 0 $V_{gate} < V_{threshold\_low}$

Basic System Components: Transistors

- NMOS



- PMOS

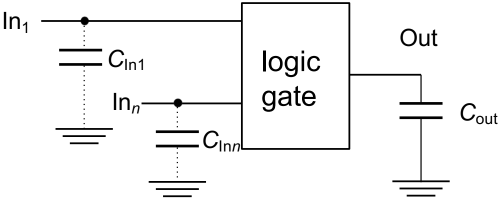


logic 1 $V_{gate} > V_{threshold\_high}$

logic 0 $V_{gate} < V_{threshold\_low}$

Logic Gates

- NMOS and PMOS transistors
  - ...implement logic gates
  - ...switch capacitances

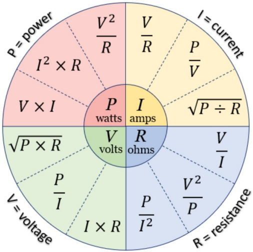


- charges move into and out of capacitors
  - input capacitances (e.g., gate capacitances)
  - output capacitances (e.g., wire length, fanout $\rightarrow$ # driven gates)

Agenda

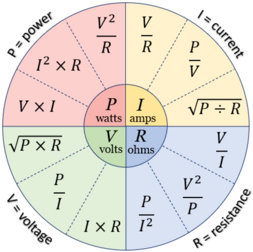
Recap: Base Units in Electric Circuits[1]

- Current I
  - flow of electric charge
  - Ampere, unit: A
- Voltage V
  - potential between two points (e.g., ground and V_{dd})
  - Volt, unit: V



[1] Digest

Recap: Base Units in Electric Circuits[1]



- Current I
  - flow of electric charge
  - Ampere, unit: A
- Voltage V
  - potential between two points (e.g., ground and V_{dd})
  - Volt, unit: V
- Power P
  - rate at which electrical energy is transferred by an electric circuit
    $\Rightarrow$ power: rate of doing work
  - Watt, unit: W $\rightarrow$ V $\cdot$ A ...or: J / s
- Energy E
  - energy that is transmitted by electricity or stored in electrical fields
    $\Rightarrow$ energy: ability to do work
  - Joule, unit: J $\rightarrow$ V $\cdot$ A $\cdot$ s ...or: W $\cdot$ s

[1] Digest

Power and Energy Demand of Systems

Definition (Energy Demand)

The energy demand E of a system is measured in joules (J) and is determined by the integral of power demand over time.

$$E_{\mathrm{op}} = \int_{t_0}^{t_1} p(t) \cdot \mathrm{d}\, t$$

Example

The energy demand E_{op} that is required to execute an operation is calculated by integrating the time function of the power demand $p(t)$ over the time $t_{op} = t_1 - t_0$ required to run the operation.

Power and Energy Demand of Systems

Definition (Power Demand)

The power demand P of a system is measured in joules per second (J/s). One joule per second equals one watt (W).

$$P_{total} = \underbrace{(C_{load} \cdot f_p \cdot A \cdot {V_{dd}}^2)}_{P_{dynamic}} + \underbrace{(I_{short} \cdot V_{dd})}_{P_{short-circuit}} + \underbrace{(I_{leak} \cdot V_{dd})}_{P_{static}}$$

Components of Power Demand

The instantaneous power demand of a circuit is split into three components: **dynamic**, **short-circuit**, and **static** power demand. Dynamic and static power demand commonly dominate.

Dynamic Power Demand

- Dynamic Power Demand
  - Capacitance $C_{load} \rightarrow$ {gate,diffusion,wire} capacitance
  - Operating Frequency $f_p \rightarrow$ clock frequency
  - Activity Factor $A \rightarrow$ fraction of clock frequency, {0...1}
  - Supply Voltage $V_{dd} \rightarrow$ (dynamic) voltage that is required for operation

$$P_{dynamic} = C_{load} \cdot f_p \cdot A \cdot {V_{dd}}^2$$

- Capacitance
- Clock Frequency
- Activity Factor
- Supply Voltage

↪ circuit design

Dynamic Power Demand

- Dynamic Power Demand
  - Capacitance $C_{load} \rightarrow$ {gate,diffusion,wire} capacitance
  - Operating Frequency $f_p \rightarrow$ clock frequency
  - Activity Factor $A \rightarrow$ fraction of clock frequency, $\{0...1\}$
  - Supply Voltage $V_{dd} \rightarrow$ (dynamic) voltage that is required for operation

$$P_{dynamic} = C_{load} \cdot f_p \cdot A \cdot {V_{dd}}^2$$

- Capacitance
  ↪ circuit design
- Clock Frequency
  ↪ limits determined by circuit
  ↪ has performance impact
- Activity Factor
- Supply Voltage

Dynamic Power Demand

- Dynamic Power Demand
  - Capacitance $C_{load} \rightarrow$ {gate,diffusion,wire} capacitance
  - Operating Frequency $f_p \rightarrow$ clock frequency
  - Activity Factor $A \rightarrow$ fraction of clock frequency, {0...1}
  - Supply Voltage $V_{dd} \rightarrow$ (dynamic) voltage that is required for operation

$$P_{dynamic} = C_{load} \cdot f_p \cdot A \cdot {V_{dd}}^2$$

- Capacitance
  ↪ circuit design
- Clock Frequency
  ↪ limits determined by circuit
  ↪ has performance impact
- Activity Factor
  ↪ depends on circuit and usage
- Supply Voltage

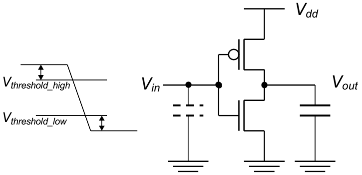
Dynamic Power Demand

- Dynamic Power Demand
  - Capacitance $C_{load} \rightarrow$ {gate,diffusion,wire} capacitance
  - Operating Frequency $f_p \rightarrow$ clock frequency
  - Activity Factor $A \rightarrow$ fraction of clock frequency, {0...1}
  - Supply Voltage $V_{dd} \rightarrow$ (dynamic) voltage that is required for operation

$$P_{dynamic} = C_{load} \cdot f_p \cdot A \cdot {V_{dd}}^2$$

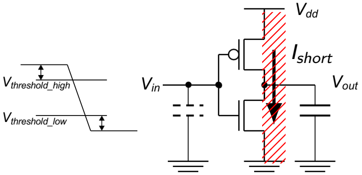
- Capacitance
  - ↪ circuit design
- Clock Frequency
  - ↪ limits determined by circuit
  - ↪ has performance impact
- Activity Factor
  - ↪ depends on circuit and usage
- Supply Voltage
  - ↪ limits determined by circuit
  - ↪ has impact on circuit switching

Short Circuit and Static Power Demand



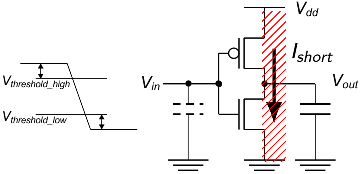
- Short-Circuit Power Demand
  - finite rise and fall times of voltages
  - NMOS/PMOS transistors conduct simultaneously $\Rightarrow P_{short} = I_{short} \cdot V_{dd}$
- Static Power Demand (Leakage)
  - gate leakage
  - sub-threshold current
  - drain junction leakage

Short Circuit and Static Power Demand



- Short-Circuit Power Demand
  - finite rise and fall times of voltages
  - NMOS/PMOS transistors conduct simultaneously $\Rightarrow P_{short} = I_{short} \cdot V_{dd}$
- Static Power Demand (Leakage)
  - gate leakage
  - sub-threshold current
  - drain junction leakage

Short Circuit and Static Power Demand



- Short-Circuit Power Demand
  - finite rise and fall times of voltages
  - NMOS/PMOS transistors conduct simultaneously $\Rightarrow P_{short} = I_{short} \cdot V_{dd}$
- Static Power Demand (Leakage)
  - gate leakage
  - sub-threshold current
  - drain junction leakage
- Trends
  - capacitances decrease $\rightarrow$ less power is required to drive the capacitance
  - lower supply voltages $\rightarrow$ lower leakage current
  - but: lower threshold voltages $\rightarrow$ higher leakage
  - gap between voltage scaling and transistor scaling results in higher power density and **dark silicon**...

Agenda

Dennard Scaling Revisited: Dark Silicon

- technology trend, state of the art
  - 2019: Core i9-7900X (14 nm, approx. 7 billion transistors, 140 W)
  - chip area unchanged ⇒ ↑ density of transistors ⇒ ↑ power density
  - result: violation of power constraints as to thermal limits
  - effect: hitting the utilization wall [8] leads to unpowered areas

Dark Silicon [2] and its impact...

Although cores fit onto die as to shrinking semiconductor scaling, they can't be powered simultaneously due to power constraints[a]

[a] at least not at with highest clock speed

Dennard Scaling Revisited: Dark Silicon

- technology trend, state of the art
  - 2019: Core i9-7900X (14 nm, approx. 7 billion transistors, 140 W)
  - chip area unchanged ⇒ ↑ density of transistors ⇒ ↑ power density
  - result: violation of power constraints as to thermal limits
  - effect: hitting the utilization wall [8] leads to unpowered areas

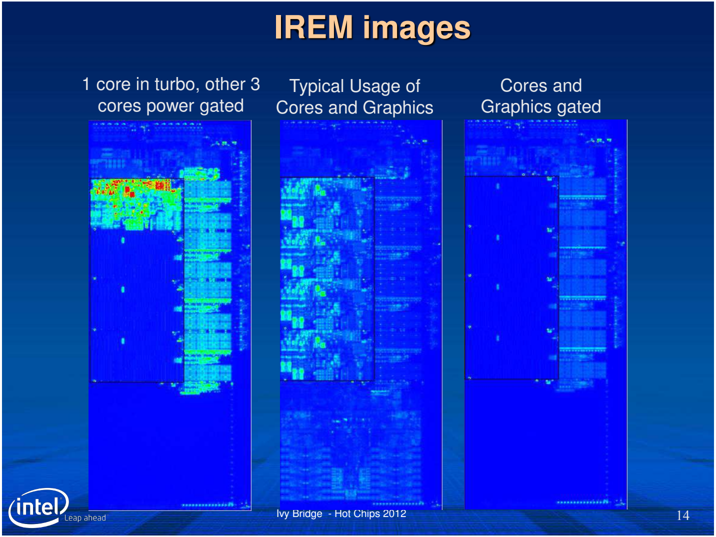
Dark Silicon [2] and its impact...

Although cores fit onto die as to shrinking semiconductor scaling, they can't be powered simultaneously due to power constraints[a]

[a] at least not at with highest clock speed

- effective (and unbeloved) counter-measures
  - switch off cores
  - run cores with reduced clock speed
  - reschedule activities

Dennard Scaling Revisited: Dark Silicon



Jahagirdar '12 [4]

Dennard Scaling Revisited: Dark Silicon

- impact of dark silicon
  - future generation systems increasingly interweave design processes of hardware and software components
  - impose challenges for operating systems
  - strict focus on **energy-awareness**

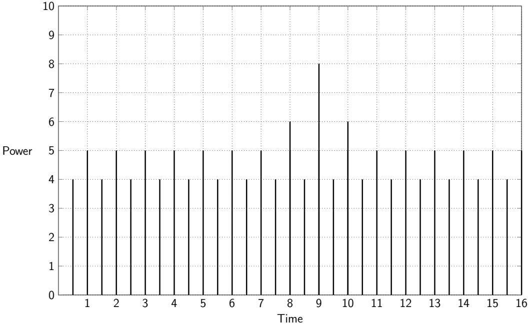
Dennard Scaling Revisited: Dark Silicon

- impact of dark silicon
  - future generation systems increasingly interweave design processes of hardware and software components
  - impose challenges for operating systems
  - strict focus on **energy-awareness**
- energy-aware system designs require...
  - comparison of systems with regards to different properties
    - power demand
    - energy demand
    - performance
    - latency
  - design criteria (static) $\rightarrow$ hardware *and* software
  - system planning (dynamic) $\rightarrow$ hardware *and* software
- **metrics** and methods for system characterization

Agenda

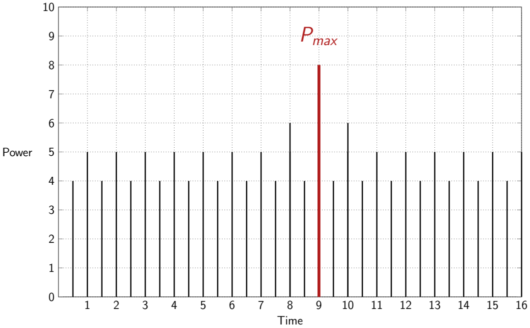
Basic Metrics: Power

- Power P (Watt, unit: W or J / s)
  - rate at which electrical energy is transferred by an electric circuit
    $\Rightarrow$ power: rate of doing work
- Power is a suitable metric for...
  - power supply constraints, cooling facilities $\rightarrow$ peak power
  - prediction of heat dissipation $\rightarrow$ average and peak power



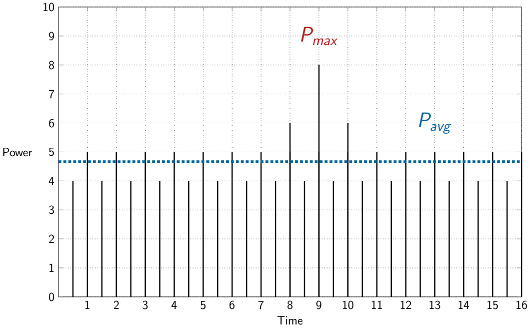
Basic Metrics: Power

- Power P (Watt, unit: W or J / s)
  - rate at which electrical energy is transferred by an electric circuit
    $\Rightarrow$ power: rate of doing work
- Power is a suitable metric for...
  - power supply constraints, cooling facilities $\rightarrow$ peak power
  - prediction of heat dissipation $\rightarrow$ average and peak power



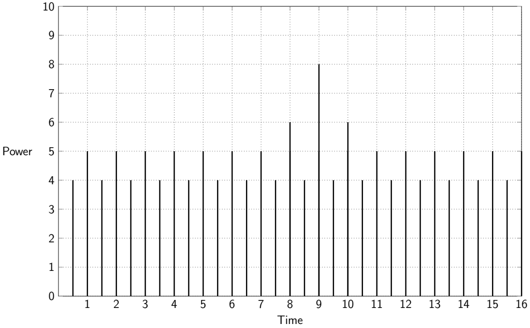
Basic Metrics: Power

- Power P (Watt, unit: W or J / s)
  - rate at which electrical energy is transferred by an electric circuit
    $\Rightarrow$ power: rate of doing work
- Power is a suitable metric for...
  - power supply constraints, cooling facilities $\rightarrow$ peak power
  - prediction of heat dissipation $\rightarrow$ average and peak power



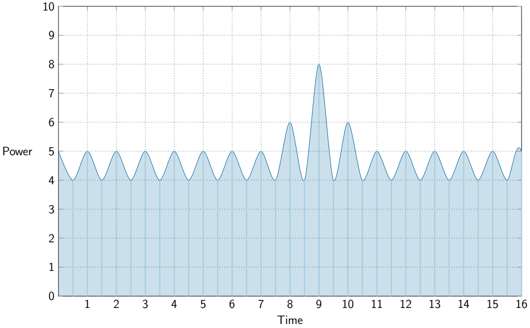
Basic Metrics: Energy

- Energy E (Joule, unit: J or $W \cdot s$)
  - energy that is transmitted by electricity or stored in electrical fields
    $\Rightarrow$ energy: ability to do work
- Energy is a suitable metric for...
  - dimensioning of electricity supplies $\rightarrow$ battery life
  - energy bill



Basic Metrics: Energy

- Energy E (Joule, unit: J or $W \cdot s$)
  - energy that is transmitted by electricity or stored in electrical fields
    $\Rightarrow$ energy: ability to do work
- Energy is a suitable metric for...
  - dimensioning of electricity supplies $\rightarrow$ battery life
  - energy bill



Basic Metrics: Power vs. Energy Revisited

- power and energy demand are insufficient metrics
- system characteristics may differ strongly even though power or energy characteristics are the same
  - performance → execution time in systems
  - latency → response time in networked systems
- extended metrics combine basic metrics (e.g., power, energy demand) with additional system properties (e.g., execution time)

*delay: time unit, i.e., measured in seconds

Basic Metrics: Power vs. Energy Revisited

- power and energy demand are insufficient metrics
- system characteristics may differ strongly even though power or energy characteristics are the same
  - performance $\rightarrow$ execution time in systems
  - latency $\rightarrow$ response time in networked systems
- extended metrics combine basic metrics (e.g., power, energy demand) with additional system properties (e.g., execution time)

- **basic** metrics are used to build different **composite** metrics
  - **energy demand** itself can be interpreted as a composite metric
  - power-delay* product (PDP):
    power demand (in Watt) $\cdot$ delay (in seconds) $\rightarrow$ energy demand (in Joule)
- more complex metrics to be explored which consider and emphasize different system properties to varying degrees...

*delay: time unit, i.e., measured in seconds

Agenda

Extended and Composite Metrics

- power-delay product (PDP): $P_{avg} \cdot t$
- energy-delay product (EDP): $E \cdot t = P_{avg} \cdot t \cdot t$
- energy-delay-squared product (ED^2P)
- energy-delay-cubed product (ED^3P)

Extended and Composite Metrics

- power-delay product (PDP): $P_{avg} \cdot t$
  - average energy consumed per switching event
  - good for fixed voltage designs
- energy-delay product (EDP): $E \cdot t = P_{avg} \cdot t \cdot t$
- energy-delay-squared product (ED^{2}P)
- energy-delay-cubed product (ED^{3}P)

Extended and Composite Metrics

- power-delay product (PDP): $P_{avg} \cdot t$
  - average energy consumed per switching event
  - good for fixed voltage designs
- energy-delay product (EDP): $E \cdot t = P_{avg} \cdot t \cdot t$
  - equal weight for changes of **energy demand** and **performance**
  - Horowitz et al. [3]
    $\hookrightarrow$ metric is misleading for systems with dynamic voltage scaling $\rightarrow$ ED^{2}P
- energy-delay-squared product (ED^{2}P)
- energy-delay-cubed product (ED^{3}P)

Extended and Composite Metrics

- power-delay product (PDP): $P_{avg} \cdot t$
  - average energy consumed per switching event
  - good for fixed voltage designs
- energy-delay product (EDP): $E \cdot t = P_{avg} \cdot t \cdot t$
  - equal weight for changes of **energy demand** and **performance**
  - Horowitz et al. [3]
    $\hookrightarrow$ metric is misleading for systems with dynamic voltage scaling $\rightarrow$ ED^2P
- energy-delay-squared product (ED^2P)
  - metric good for fixed micro architecture with dynamic voltage scaling
  - Brooks et al. [1]
- energy-delay-cubed product (ED^3P)
  - further emphasize on performance, used for high-performance scenarios
  - Srinivasan et al. [6]

Subject Matter

- **power** and **utilization walls** (dark silicon) forces drastic redesign of computing systems for energy awareness
- energy demand of computing systems must be seen in due **consideration** of other **non-functional properties** (e.g., performance)
- available **metrics** must be suitable for individual use

- reading list for Lecture 3:
  - Vivek Tiwari et al.
    Power Analysis of Embedded Software: A First Step Towards Software Power Minimization
    IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 1994.

Reference List I

[1] BROOKS, D. M. ; BOSE, P. ; SCHUSTER, S. E. ; JACOBSON, H. ; KUDVA, P. N. ; BUYUKTOSUNOGLU, A. ; WELLMAN, J. ; ZYUBAN, V. ; GUPTA, M. ; COOK, P. W.:
Power-aware microarchitecture: design and modeling challenges for next-generation microprocessors.
In: *IEEE Micro* 20 (2000), Nov, Nr. 6, S. 26–44

[2] ESMAEILZADEH, H. ; BLEM, E. ; AMANT, R. S. ; SANKARALINGAM, K. ; BURGER, D. :
Dark silicon and the end of multicore scaling.
In: *Proceedings of the 38th Annual International Symposium on Computer Architecture (ISCA)*, 2011, S. 365–376

[3] HOROWITZ, M. ; INDERMAUR, T. ; GONZALEZ, R. :
Low-power digital design.
In: *Proceedings of 1994 IEEE Symposium on Low Power Electronics*, 1994, S. 8–11

[4] JAHAGIRDAR, S. ; GEORGE, V. ; SODHI, I. ; WELLS, R. :
Power management of the third generation Intel Core micro architecture formerly codenamed Ivy Bridge.
In: *Proceedings of the IEEE Hot Chips 24 Symposium (HCS)*, 2012, S. 1–49

Reference List II

[5] Pollack, F. J.:
New microarchitecture challenges in the coming generations of CMOS process technologies.
In: *Proceedings of the 32nd Annual ACM/IEEE International Symposium on Microarchitecture*, 1999

[6] Srinivasan, V. ; Brooks, D. ; Gschwind, M. ; Bose, P. ; Zyuban, V. ; Strenski, P. N. ; Emma, P. G.:
Optimizing pipelines for power and performance.
In: *Proceedings of the 35th Annual IEEE/ACM International Symposium on Microarchitecture*, 2002, S. 333–344

[7] Sutter, H. :
The free lunch is over: A fundamental turn toward concurrency in software.
In: *Dr. Dobb's journal* 30 (2005), Nr. 3, S. 202–210

[8] Venkatesh, G. ; Sampson, J. ; Goulding, N. ; Garcia, S. ; Bryksin, V. ; Lugo-Martinez, J. ; Swanson, S. ; Taylor, M. B.:
Conservation Cores: Reducing the energy of mature computations.
In: *Proceedings of the 15th International Conference on Architectural Support for Programming Languages and Operating Systems*, 2010, S. 205–218