

Energy Efficiency in Computing (1)

CERN Academic Training – May 2016

Andrzej Nowak

http://tik.services

technology innovation knowledge

TIK Services





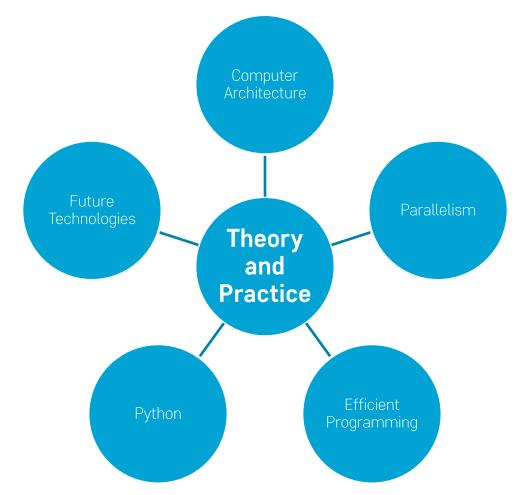


Technology

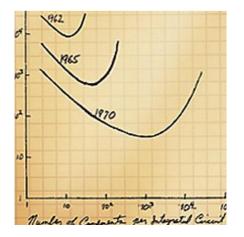
Innovation

Knowledge

TIK Training



Outline



Day 1: Silicon, hardware

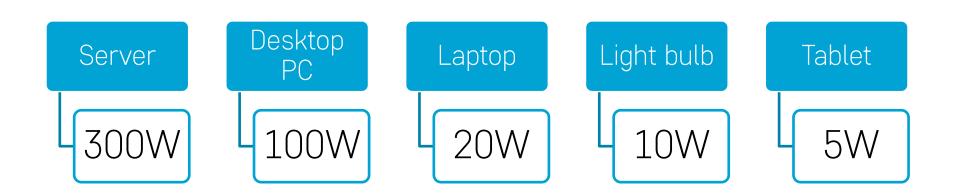
Day 2: Datacenters, software, future technologies

| 🅘 Analysis Target 🛛 🏯 Analys | is Type | M Summary | • | CPU Sleep St | stes 🙆 C | ore Wake |
|------------------------------|--------------|---------------|----------|--------------|-----------|----------|
| n / Wake-up Object / Core | | | | | | |
| Wake-up Object / Core | Total | Wake-up Count | • | Process Name | ProcessID | Thread |
| | | 5 | 68 | socwatch | 6719 | 6719 |
| | | 5 | 48 | irq/67-intel | 78 | 78 |
| | | 2 | 09 | kworker/0:8 | 1111 | 1111 |
| | | 1 | 95 | mediaserver | 151 | 270 |
| | | 1 | 60 | mediaserver | 151 | 270 |
| | | | 35 | mediaserver | 151 | 270 |
| | | 1 | 93 | kworker/0:4 | 23310 | 23310 |
| Selected 1 row(s): | | | 86 95 | kworker/0:3 | 6560 | 6560 |
| | 4 | | | | | |
| 21.369; Ss 22s | i di li A | 22.55 | 23 | 23.5 | | Core |
| Wake-up Object | 82.701 m | s Inline Mo | - | | Functions | V V |

Basics

1 Watt = 1 Joule / second (power = energy / second)

Everyday devices



Practical considerations

How much does it cost to charge the iPhone 6 and iPhone 6 Plus?



iPhone 6 \$0.47/year (3.8 kWh)



iPhone 6 Plus \$0.52/year (4.2 kWh)

Opower 2014

Image: opower

Practical considerations

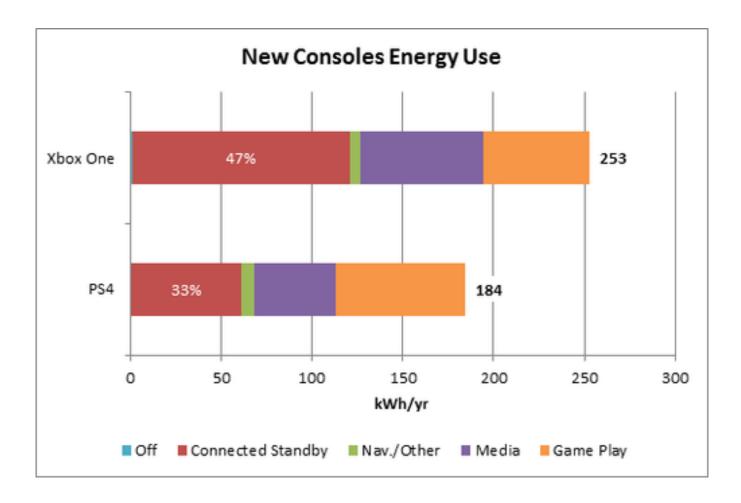
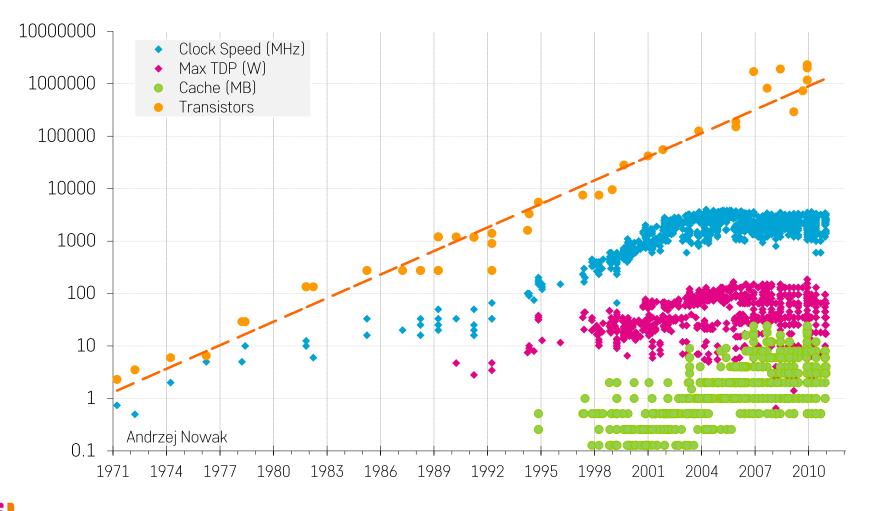


Image: NRDS

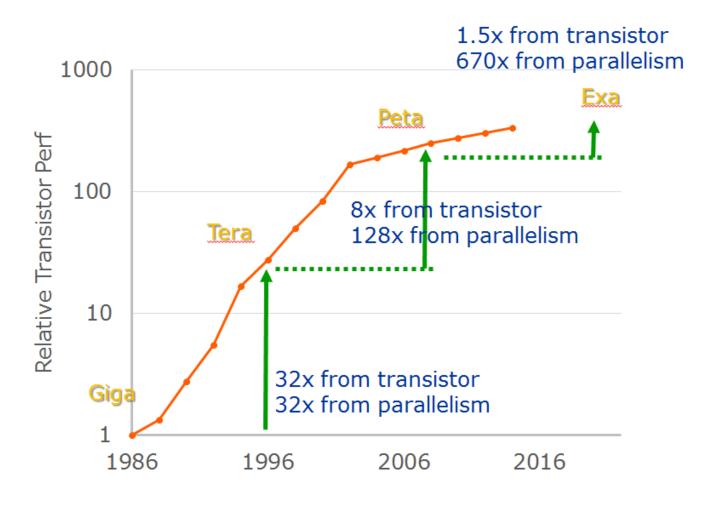
Energy-efficient Silicon



Moore's Law and power Since 1971 – based on 1'700 Intel CPUs



Gain (and Pain)

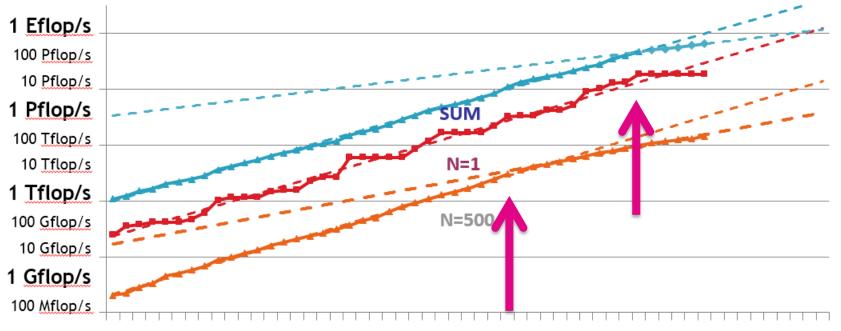


IK. Energy Efficiency in Computing (1/2)

Image: S. Borkar/Intel

Moore's Law and Top500

PROJECTED PERFORMANCE DEVELOPMENT

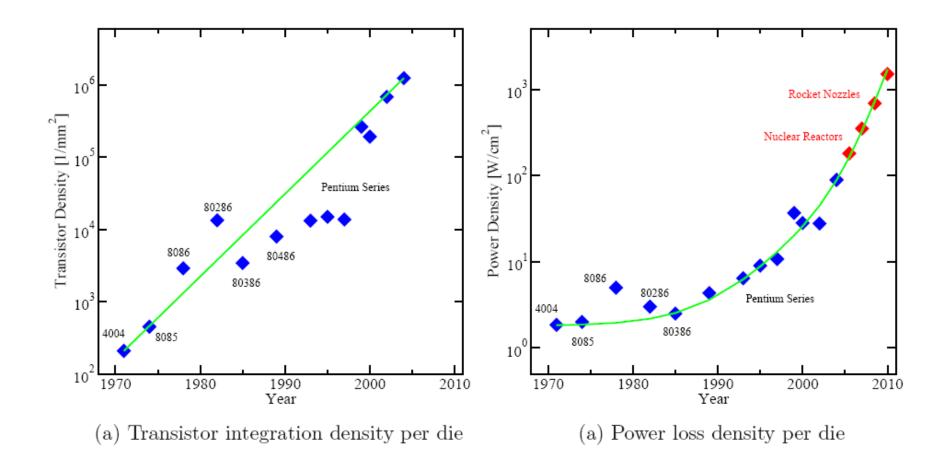


1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018 2020

Image: E. Strohmaier / HPC Wire

14

The power problem



Energy Efficiency in Computing (1/2)

Image: TU Wien

Transistor operation

Operating Regions

Revisit transistor operating regions

| Region | nMOS | pMOS |
|--------|------------|------------|
| A | Cutoff | Linear |
| В | Saturation | Linear |
| С | Saturation | Saturation |
| D | Linear | Saturation |
| E | Linear | Cutoff |

Image: Internet

Transistor power

$P = ACV^2f + tAVI_{short}f + VI_{leak}$

$$f_{max} = \frac{(V - V_{threshold})^2}{V}$$

Energy Efficiency in Computing (1/2)

By NYU 18-May-16

Transistor power facts

Dynamic power scales with the square of the voltage

• Need to reduce V as much as possible

Transistors have a gate threshold voltage for switching from off to on

• Can't reduce V too much

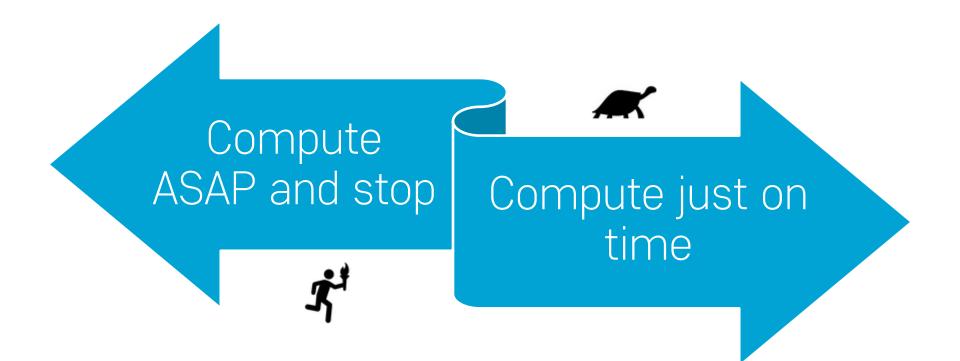
Switching speed ~inversely proportional to V

• Can't reduce V too much

Transistor power problems

- Too much power lost through "high" voltage
 Dynamic voltage scaling
- Power consumption vs frequency goes with the power law
 - Reduce frequency statically
 - Reduce frequency dynamically
- Switching delays. Solutions:
 - Make wires shorter
 - Make gates smaller
 - Use faster materials

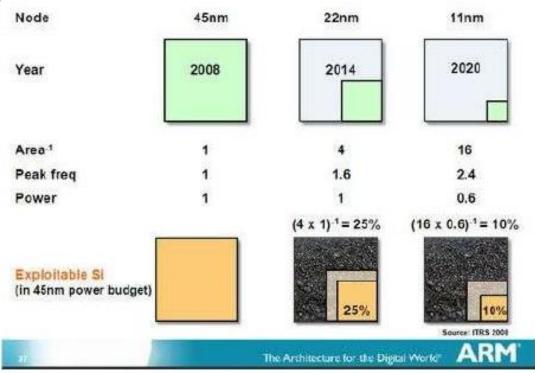
Computation strategies



20

Dark Silicon

- Systems increasingly limited by power consumption, not number of transistors
- • "Dark Silicon": Most of the chip will be OFF to meet thermal limits



Alleviating the problem

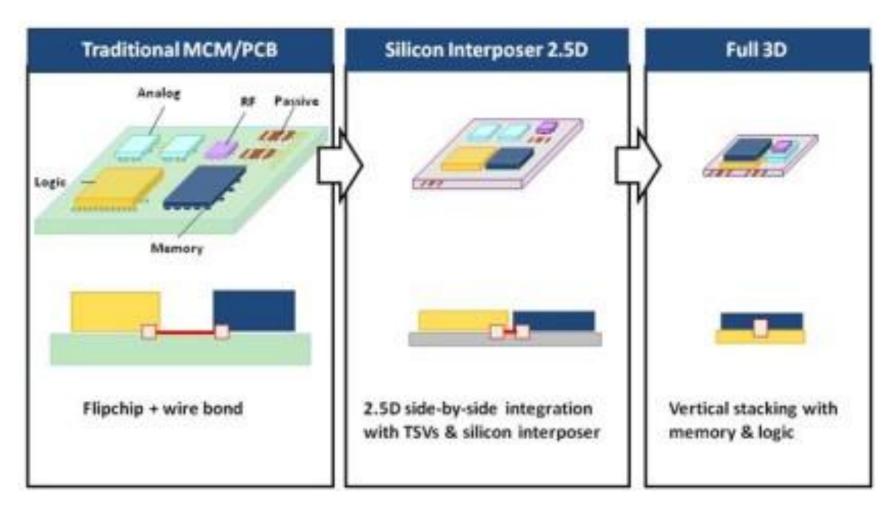
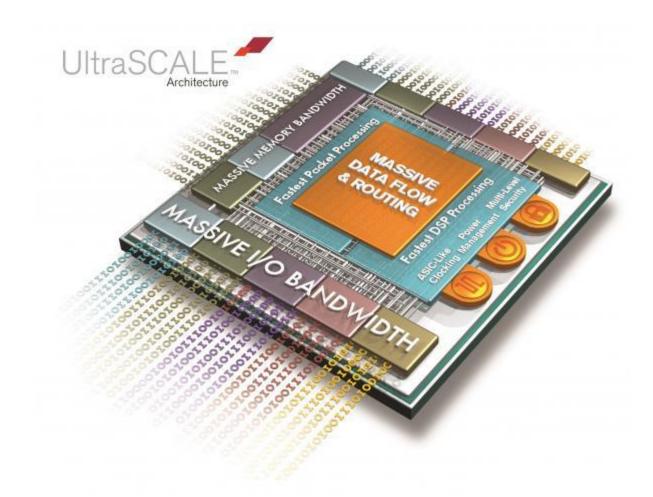


Image: EE Times





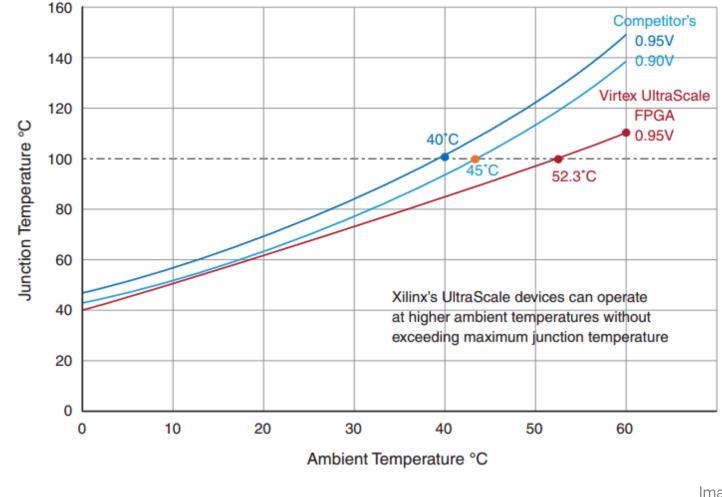
Reconfigurable logic FPGAs



Reconfigurable logic Power constraints as seen by Xilinx

| A | | C D E F | G H I | J K | L M N O P |
|-----|----------------------|------------------------------|------------------------|----------------|-------------------------------------------|
| 1 | S XILII | | ۱x Power Estimator () | (PE) - 2016.1 | ۹ ۴ ۵ |
| 2 | | NA® Kinte | ex® UltraScale™, Virte | ex® UltraScale | Release: 13-Apr-2016 |
| 3 | Import File | Export File Quick Estimation | te Manage IP | E Snapshot | Set Default Rates |
| 5 | Project | | | | |
| 6 _ | Settings | | | Summai | ry |
| 7 | [| Device | Total On-Chip Power | 0.5 W | 0% • Transcelver 0.000W |
| 3 | Family | Kintex UltraScale | | 0.5 W | 0% • I/O 0.000W |
| | Device | XCKU040 | Junction Temperature | 25.8 °C | 0% • Core Dynamic 0.000W |
| 0 | Package | FBVA900 | Thermal Margin | 74.2°C 39.6W | 100% • Device Static 0.458W |
| 1 | Speed Grade | -1L (0.9V) | Effective ΘJA | 1.8 °C/W | Power supplied to off-chip devices 0.000W |
| 2 | Temp Grade | Industrial | | | |
| 3 | Process | Typical | — On-Chip Power | | Power Supply |
| 4 | Voltage ID Used | | Resource | Power | Source Voltage Total (A |
| 5 | Characterization | Production (± 20% accuracy) | (Jump to sheet) | (VV) (%) | V _{CCINT} 0.900 0.135 |
| 6 | | | CLOCK | 0.000 0 | V _{CCINT_IO} 0.900 0.014 |
| 7 | En | vironment | LOGIC | 0.000 0 | V _{CCBRAM} 0.950 0.011 |
| 8 | Junction Temperature | 🗆 User Override | BRAM | 0.000 0 | V _{CCAUX} 1.800 0.096 |
| 9 | Ambient Temp | 25.0 °C | DSP | 0.000 0 | V _{CCAUX_IO} 1.800 0.065 |
| | Summary | Snapshot Graphs IP_Manage | er Clock Logic IO B | RAM DSP CLKN | MGR GTH Other User Relea |

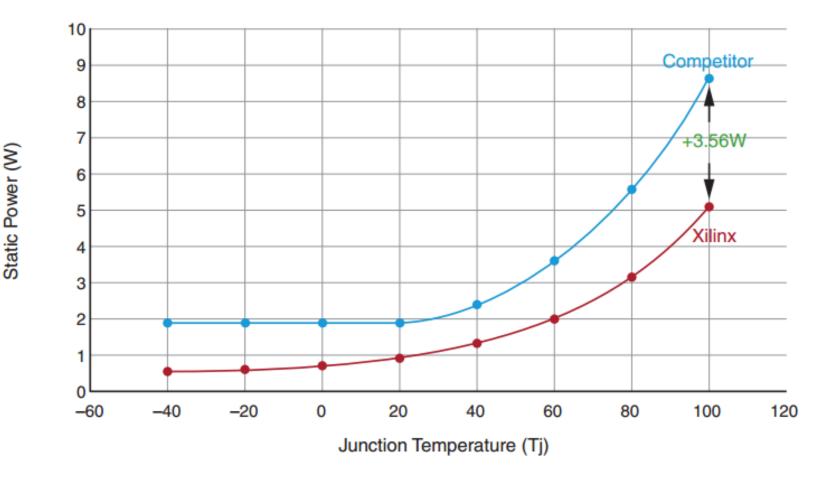
Reconfigurable logic Environmental constraints as seen by Xilinx



tik. Energy Efficiency in Computing (1/2)

Image: Xilinx

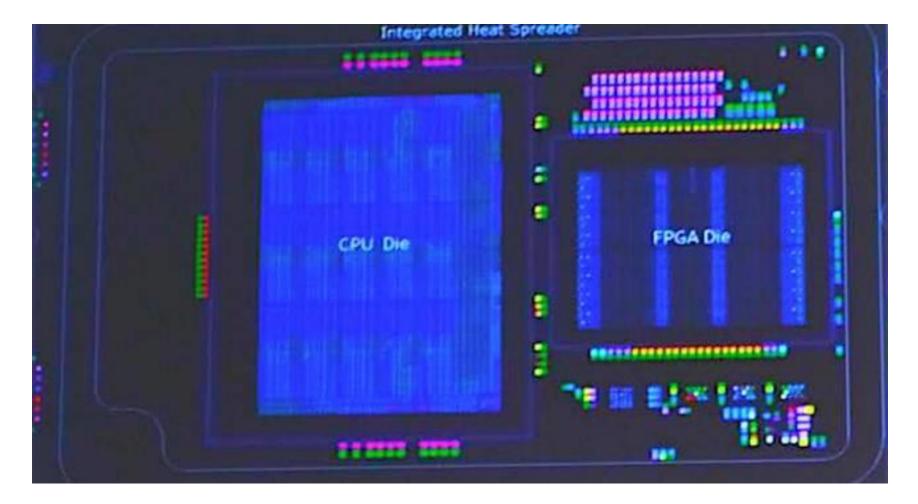
Reconfigurable logic Environmental constraints as seen by Xilinx



tik. Energy Efficiency in Computing (1/2)

Image: Xilinx

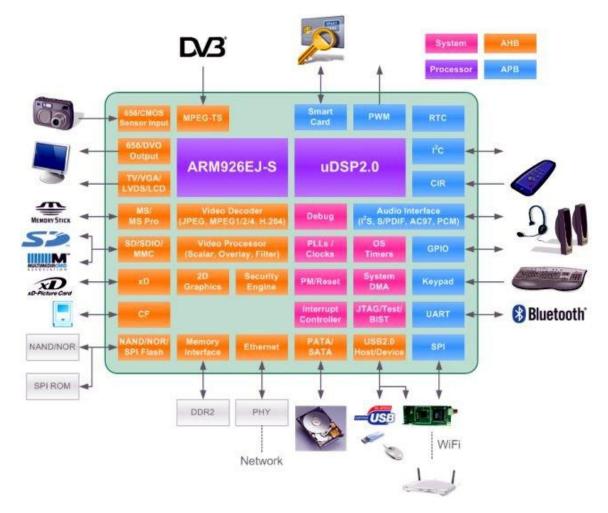
Reconfigurable logic



tik. Energy Efficiency in Computing (1/2)

Image: The Next Platform

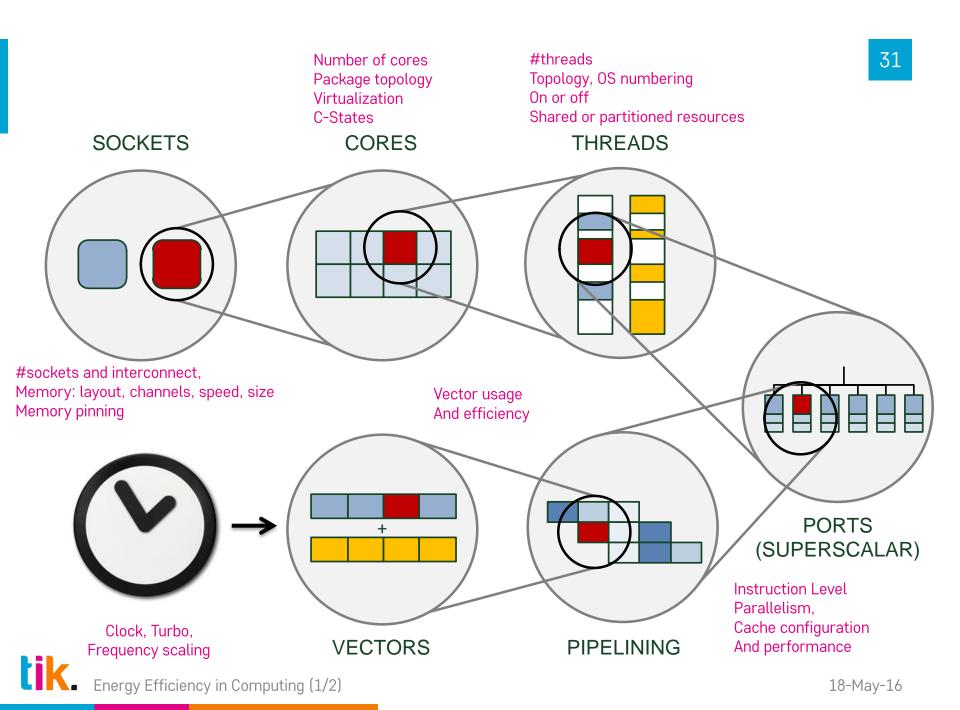
SoC



Rough CPU energy breakdown

- Clock distribution ~10%
- L1/L2 ~20%
- Exec ~35-40%
- Routing/Links ~25%
- Other ~10%

30



CPU C-States Idle power and latency

Table 1: CPU Idle Power States in Nexus 4.

| Idle | Name | Idle System | Latency |
|------------------------------|---------------------------|-------------|---------------------|
| State | | Power (mW) | $(\mu S)^{\dagger}$ |
| C0 | Wait for Interrupt | 433 | 1 |
| C1 | Retention | 390 | 415 |
| C2 | Power Collapse Standalone | 330 | 1300 |
| C3 | Power Collapse | 200 | 2000 |
| Without entering idle states | | 1,060 | 0 |

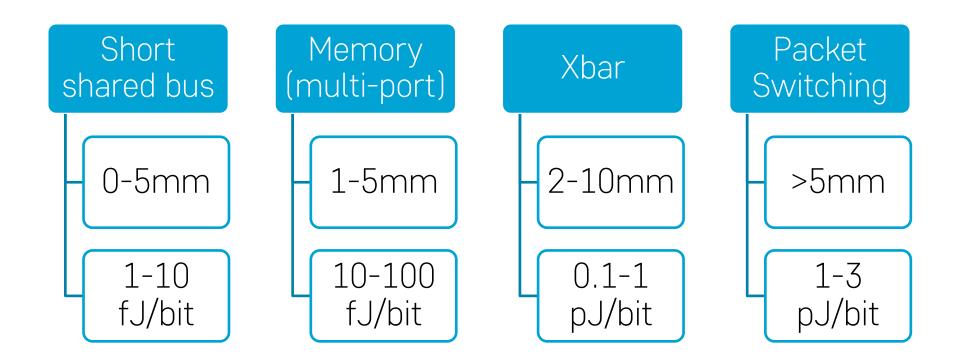
†: The data is obtained from the Nexus 4 kernel source code.

Image: Microsoft

Energy per operation

- Theoretical FLOPS
 - Intel Westmere: 1.7nJ/flop
 - Intel Haswell: 110pJ/flop
 - NVIDIA Fermi: 225pJ/flop
 - ARM (Cortex-A7): 90-150pJ/flop
 - ARM (Cortex-A15): 250-1200pJ/flop
 - Exascale target for ops: 20pJ/flop (= Exa @20MW)
- Communication :
 - Core-to-core: ~10pJ/byte
 - Chip-mem: ~150pJ/byte
 - Chip-chip: ~100pJ/byte

Interconnect energy

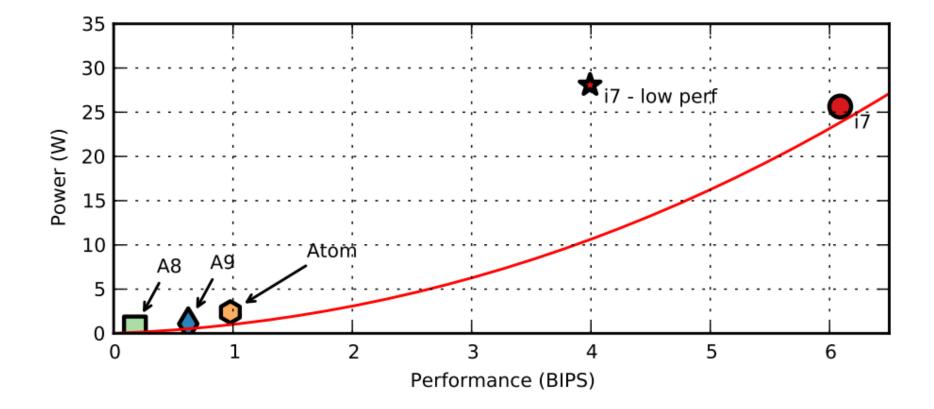


34

Storage energy cost

| | SRAM | DRAM | FLASH | Disk |
|------------|----------|------------|------------|-----------|
| Energy/bit | 1 pJ/bit | 100 pJ/bit | 500 pJ/bit | 100nJ/bit |

Intel Architecture vs. ARM Power



HPCA 2013

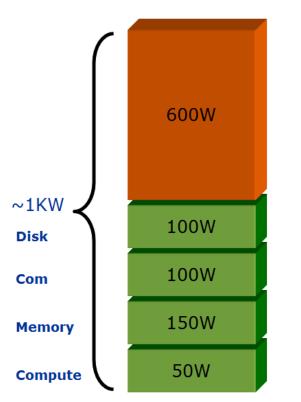
RAM energy cost All about density

| Memory Module | Capacity (GB) | Метогу Туре | Power Consumption (W) | Power Consumption (W/GB) |
|----------------|---------------|--------------------------------------------------------------------------------------|-----------------------|--------------------------|
| AL48M72F4GKF8S | 16 | DDR3, Registered, ECC, 4 rank | 8.710 | 0.54 |
| AL24M72E4BKH9S | 8 | DDR3, Registered, ECC, 2 rank | 6.132 | 0.77 |
| AL12M72B8BKH9S | 4 | DDR3, Registered, ECC, 2 rank, memory chips with 256Mx8 organization (2-Gb chips) | 2.934 | 0.73 |
| AL56M72B8BJH9S | 2 | DDR3, Registered, ECC, 2 rank, memory chips with 128Mx8 organization (1-Gb chips) | 5.132 | 2.57 |
| AL28M72A8BJH9S | 1 | DDR3, Registered, ECC, 1 rank | 2.241 | 2.24 |
| AQ12M72E8BKH9S | 4 | DDR3, Unbuffered, ECC, 2 rank, memory chips with 256Mx8 organization (2-Gb chips) | 2.214 | 0.55 |
| A028M72D8BJH9S | 1 | DDR3, Unbuffered, ECC, 1 rank | 1.387 | 1.39 |
| J28K72F8BJE6S | 1 | DDR2, Unbuffered, ECC | 1.872 | 1.87 |
| AP56K72G4BHE6S | 2 | DDR2, FB-DIMM, ECC | 13.683 | 6.84 |

Image: ATP

Energy spending in a server

Teraflop system today



Decode and control Address translations... Power supply losses **Bloated with inefficient architectural features**

10TB disk @ 1TB/disk @10W

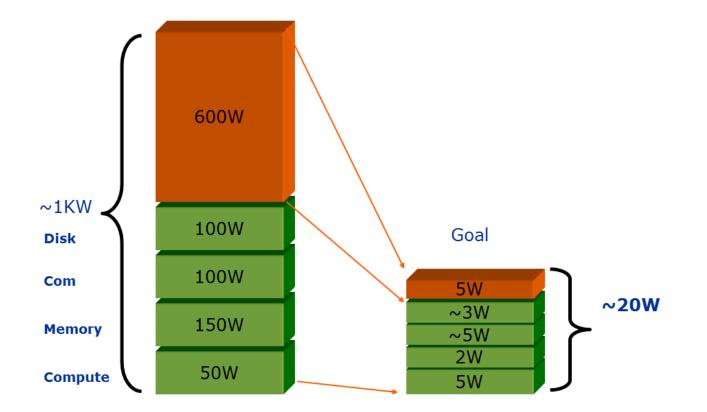
100pJ com per FLOP

0.1B/FLOP @ 1.5nJ per Byte

50pJ per FLOP

From S. Borkar / Intel

Energy spending in a server



tik. Energy Efficiency in Computing (1/2)

From S. Borkar / Intel

Thank you

e-mail: an@tik.services

http://tik.services





All content which is original in this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.