CS 4230 : Parallel Programming

Ganesh Gopalakrishnan, Mohammed S. Mahfoudh and Simone Atzeni
School of Computing
University of Utah

URL: www.eng.utah.edu/~cs4230

Lecture 2: The Physics of Computing

- ***Why do we care what the hardware does?***
  - **It used to be possible to abstract away hardware**
    - **When computing began, one had to have spare Tubes / Transistors, HW debugging aids, hammers…**
    - **Soon the hardware became stable and predictable and everyone forgot about the hardware.**
    - **Especially when the cost model was based on the number of instructions executed per second**
    - **Energy was an "afterthought"**
  - **We are now in a strange new world**
    - **Parallel activities running on heterogeneous cores deliver the overall performance**
      - **measured in terms of runtime as well as energy consumption**
    - **Hardware has begun flaking as well**
  - **Many programming notations that also address heterogeneous memory types**
- ***A programmer wanting the maximum computing efficiency has to be aware of the nature of hardware***
  - **Things are also changing so rapidly**
    - **So to keep abreast, we need to know what today's hardware organizations are, and their characteristics**
- ***MUCH IS NOW CENTERED AROUND ENERGY, LOCALITY OF DATA ACCESSES, REPRODUCIBILITY OF RESULTS, etc.***

Some Basics about Electricity

- **Since computers run on electricity, some basic EE must be appreciated and actually enjoyed**

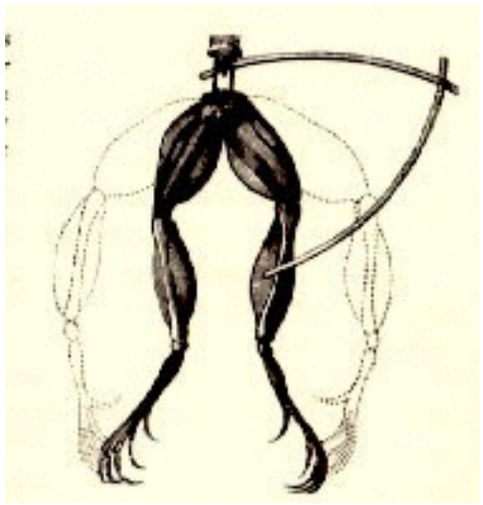
- **So let's start with**

  - **Voltage, Current, Resistance, Inductance, Capacitance**

Galvanic Frog and Alessandro Volta

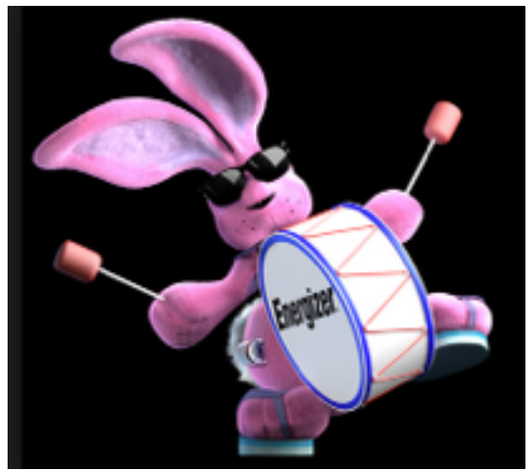
- **Voltage aka Potential Difference**
  - **Similar to column height of water in a jar**

Luigi Galvani 1737-1798



Pet Subject Kermit the Twitcher ?-?

Alessandro Volta 1745-1827



Chip Voltages

- 5v, 3.3v, 1.2v, (very little room any more)
- Reason : Transistors have threshold voltages
- Reason : Noise margins are needed
- The lowest energy one can “get away with” is really limited by noise
  - Need to store state reliably
  - Need to communicate reliably

R, C, L, V, I

- ***Current :***
  - **Similar to water flow**
- ***Resistance***
  - **Similar to water resistance (in a pipe) : V = I . R (Ohm's Law)**
- ***Capacitance***
  - **Similar to storage ability**
    - **An ideal battery of 6V ~ an infinite ocean of 6" ht**
      - **As you draw more, the level stays 6"**
    - **A capacitor holding 6V ~ a finite vessel of 6" ht**
      - **As you draw more, the level ("head") reduces**
        - **The lower the head, the lower the pressure, hence Exp decay**
  - **If you look at the "current" coming out of a finite vessel (water flow rate), it is proportional to the voltage difference**
  - **I = C dv / dt**
    - **Take a VERY LARGE finite capacity vessel : Now if you create pulsating pressure at its bottom, the head height hardly changes in response —> you can absorb the pressure pulsation —> Capacitor appears a "dead short"**
- ***Inductance***
  - **Similar to inertia**
  - **Try moving a mass suddenly : you can't ; set it in motion : can't stop suddenly (huge force needed)**
  - **V = L di / dt because if you take an L (inductor) and cause a high di / dt (sudden change in motion), you get a huge**

Typical Capacitances in Chips and Elsewhere

- ***Gate input capacitance of inverter***
  - **https://inst.eecs.berkeley.edu/~ee42/fa01/LectNotes/42_24.pdf**
  - **Typically measured in Femto Farads (e.g. 1fF)**
  - **Which is 10^{-15} farads**
- ***Capacitance of a Super Capacitor :***
  - **Can buy a 5000 Farad Supercap!**
    - **Can jump-start your car off of a 330 F cap array!**
      - **http://www.youtube.com/watch?v=Pk869dPTxNg**
    - **Used in regenerative braking, cordless screw-drivers (new rapid-recharge rapid-discharge versions) — see wikipedia**

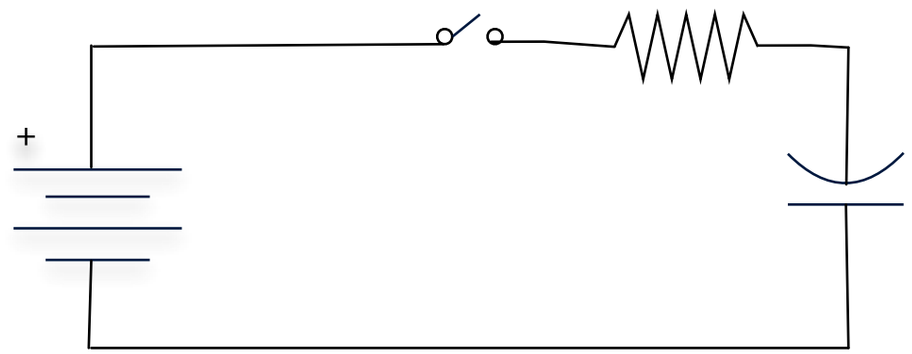
Typical Capacitances in Chips and Elsewhere

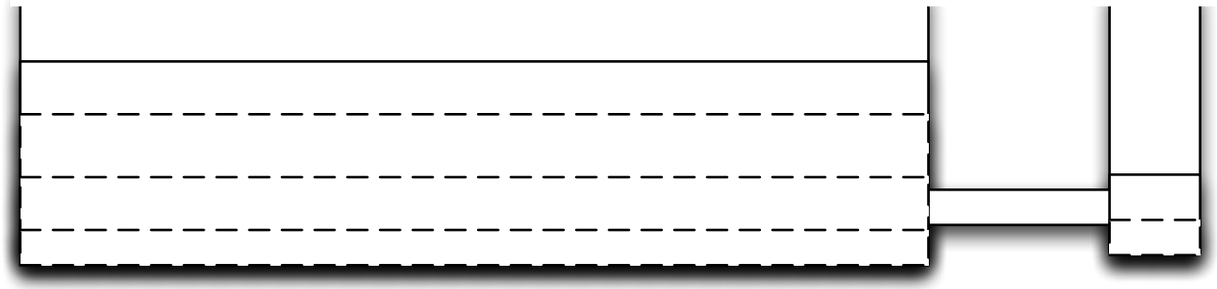
- ***Things to remember about Capacitances and Chips***
  - **A lot of the power in chips is spent charging and discharging the gate capacitors**
  - **If you have 1 transistors and have 1fF per gate, the total gate capacitance presented is 1 .. of course 1 fF**
    - **At 2 GHz, 1 fF looks like an open circuit**
    - **Impedance of a cap = 1/j.omega.C = -j / omega . C**
    - **Omega at 2 GHz = 2.pi.f = 6.28 * 2 * 10^9**
    - **C = 1 fF = 10^-15**
    - **Impedance = -j / (10^10 * 10^-15) = 100 kilo ohms (forget -j)**
    -

Typical Capacitances in Chips and Elsewhere

- ***Things to remember about Capacitances and Chips***
  - **A lot of the power in chips is spent charging and discharging the gate capacitors**
  - **If you have 1B transistors and have 1fF per gate, the total gate capacitance presented is 1 micro Farads — pretty high.**
    - **At 2 GHz, 1 micro Farads looks like a dead short.**
    - **Impedance of a cap = 1/j.omega.C = -j / omega . C**
    - **Omega at 2 GHz = 2.pi.f = 6.28 * 2 * 10^9**
    - **C = 1 micro F = 10^-6**
    - **Impedance = -j / (10^10 * 10^-6) = 100 micro ohms (forget -j)**
      - **You get the idea… little capacitances add up fast!**

Here is an absolutely delightful analogy… and a puzzle





- The above situations are quite similar
- If you charge up the Cap, the amount of charge in it is CV
- The energy within it is 0.5 C V^2 Joules
- The battery moved a charge of CV at voltage V
- Thus the battery expended C V^2 Joules
- Where did half the energy go?
- This is independent of R !!
- This means that if you fill a little vessel (a cap) from a huge one (a battery), then
  - Even if you spent a certain amount of energy pushing the water thru R (narrow orifice)
  - The amount of energy present in the smaller vessel is always half of what you spent
  - The other half goes out as heat
  - Can’t avoid this!
  - See the derivation at http://hyperphysics.phy-astr.gsu.edu/hbase/electric/capeng2.html

Typical Capacitances in Chips and Elsewhere

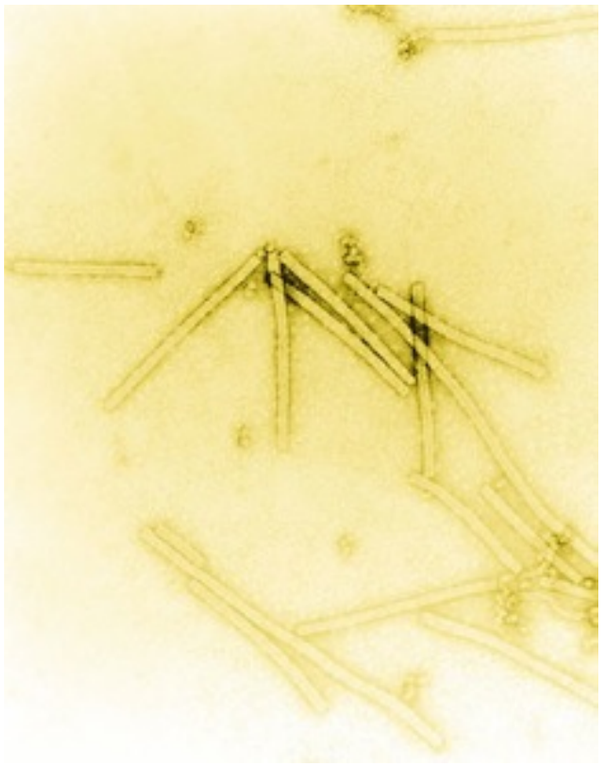
- ***Things to remember about Capacitances and Chips***
  - **If you stay within a chip, a very tiny gate/transistor has to drive another tiny gate/transistor**
    - **This is energy efficient**
  - **The moment you have to cross chips, you have to amplify the tiny gate's output a HUGE amount**
  - **Then drive the pins**
  - **Then get into the other chip via a pad**
    - **(PADs have "lightning arrestors" and such..)**
  - **Then drive the other itty-bitty gate of a chip**
    - **All makes cross-chip communication VERY inefficient!**

How big are today's transistors wrt a virus ?

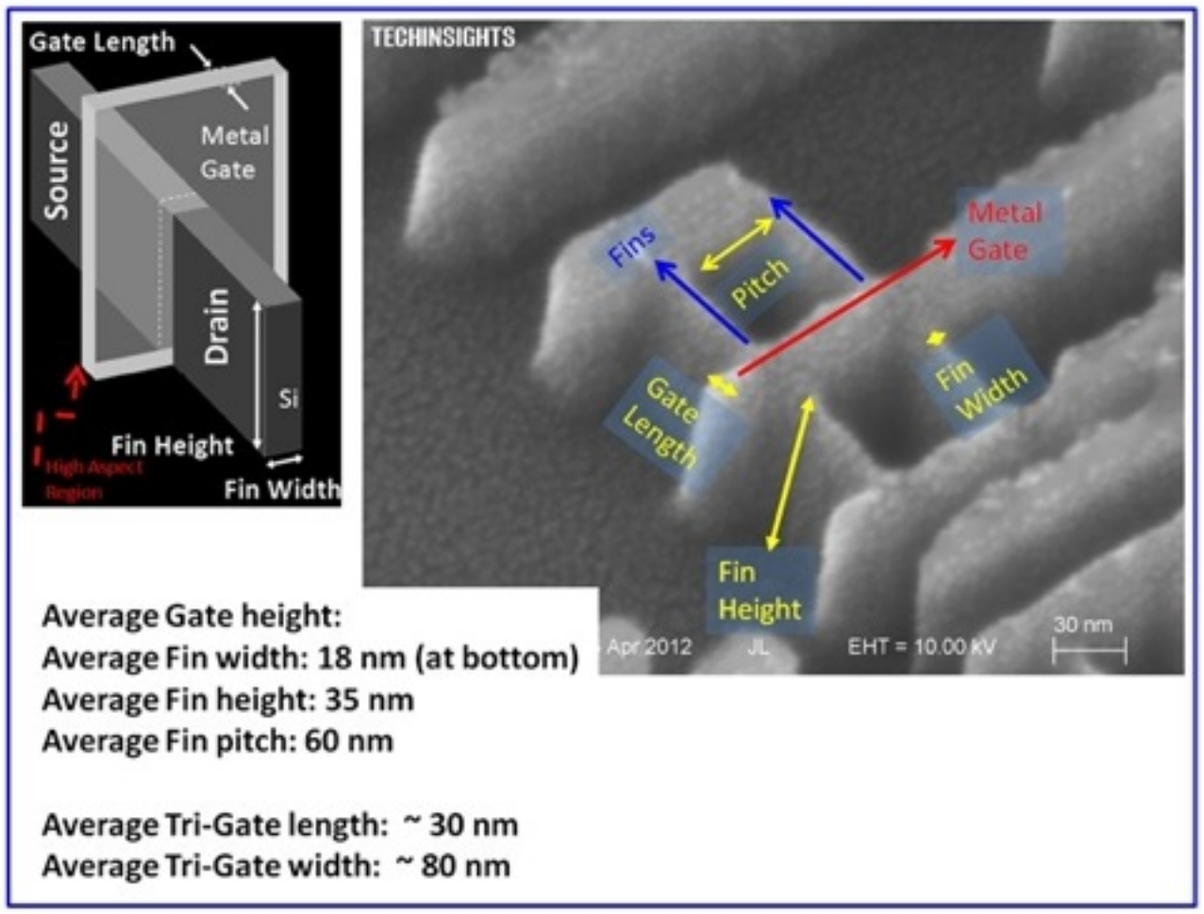
Virions are ~300 nm in length and ~18 nm in diameter.[12]

http://en.wikipedia.org/wiki/Tobacco_mosaic_virus

https://www.google.com/search?q=cmos+22nm+transistor+photo&espv=2&tbm=isch&imgil=Sw0jVD4bePQvtM%253A%253B5_zKVJSGKEGQbM%253Bhttp%25253A%25252F%25252Fwww.eetimes.com%25252Fdocument.asp%25253Fdoc_id%2525253D1281328&source=iu&usg=__a6S4ge1fsS3Vs60Mjbpv7a5S3HU%3D&sa=X&ei=RUveU6-zDcGmyASs5oCIDA&ved=0CCsQ9QEwBg&biw=1440&bih=713#facrc=_&imgdii=_&imgrc=Sw0jVD4bePQvtM%253A%3B5_zKVJSGKEGQbM%3Bhttp%253A%252F%252Fm.eet.com%252Fmedia%252F1169841%252F120906_intel_22_1.jpg%3Bhttp%253A%252F%252Fwww.eetimes.com%252Fdocument.asp%253Fdoc_id%253D1281328%3B628%3B475



heh heh, I can wear > 10 transistors head to toe !!



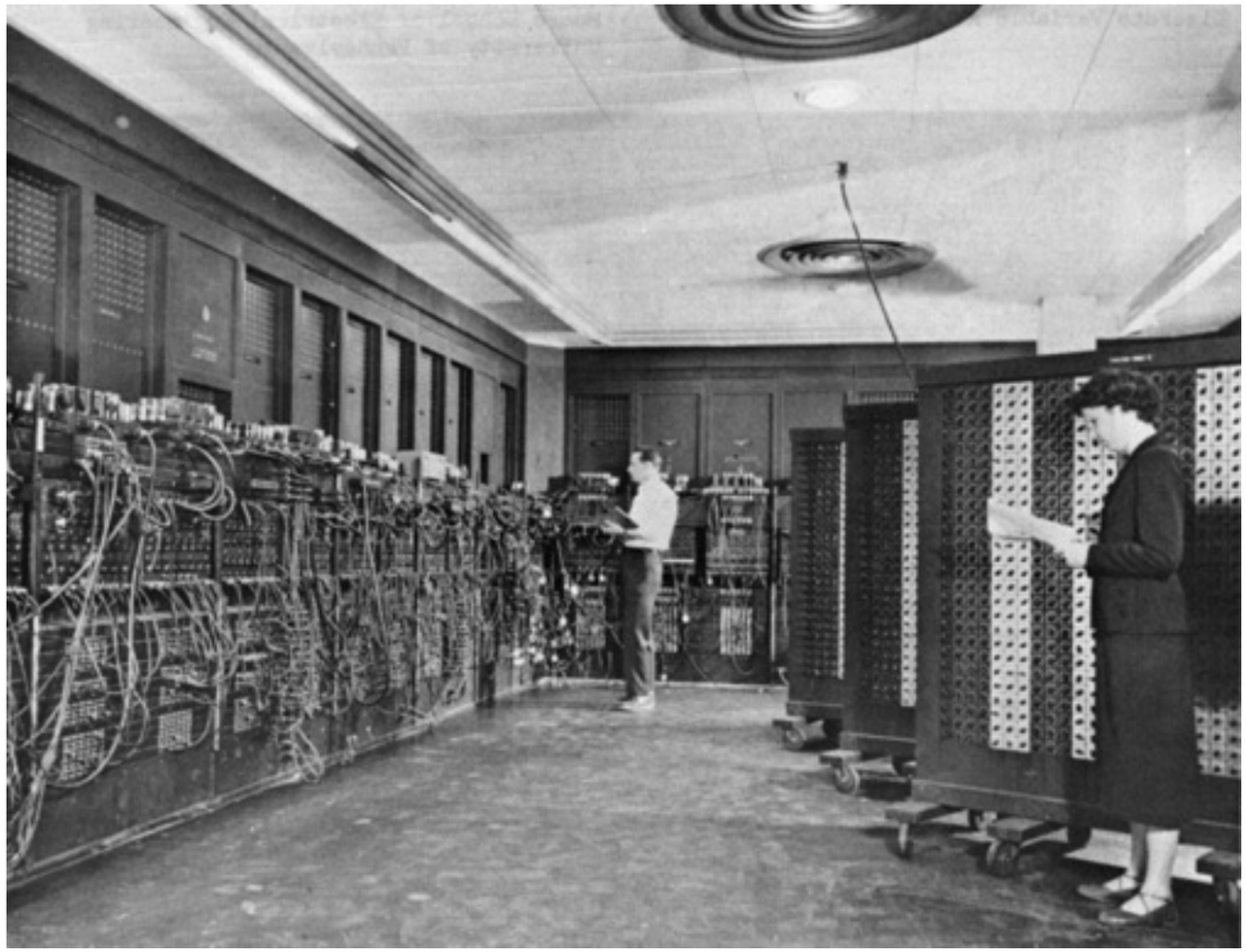
Lab Experiments Planned

- ***Saliva Ring Oscillator***
  - **Demo that we can put 5 inverters in a loop, and power the whole circuit using a suitable galvanic cell**
    - **e.g. paper clip, copper wire, tongue**
      - **drink sprite and repeat, if it does not work :)**
- ***Power CMOS operated by touch***
  - **Cue from Sutherland's experiment**
    - **http://web.cecs.pdx.edu/~mroncken/ARC-2012-is15_keynote_Vail.pdf**
    - **Car brake-light turned off-on by touching gate by one finger and the other finger on +ve**
      - **Light turns on — and stays on (gate cap holds charge)**
      - **Hey, I reproduced this - here is a Youtube Video of me doing this experiment (with a lot of new twists !!)**
    - **When -ve terminal is touched and gate is touched, the gate can be discharged —> light off**

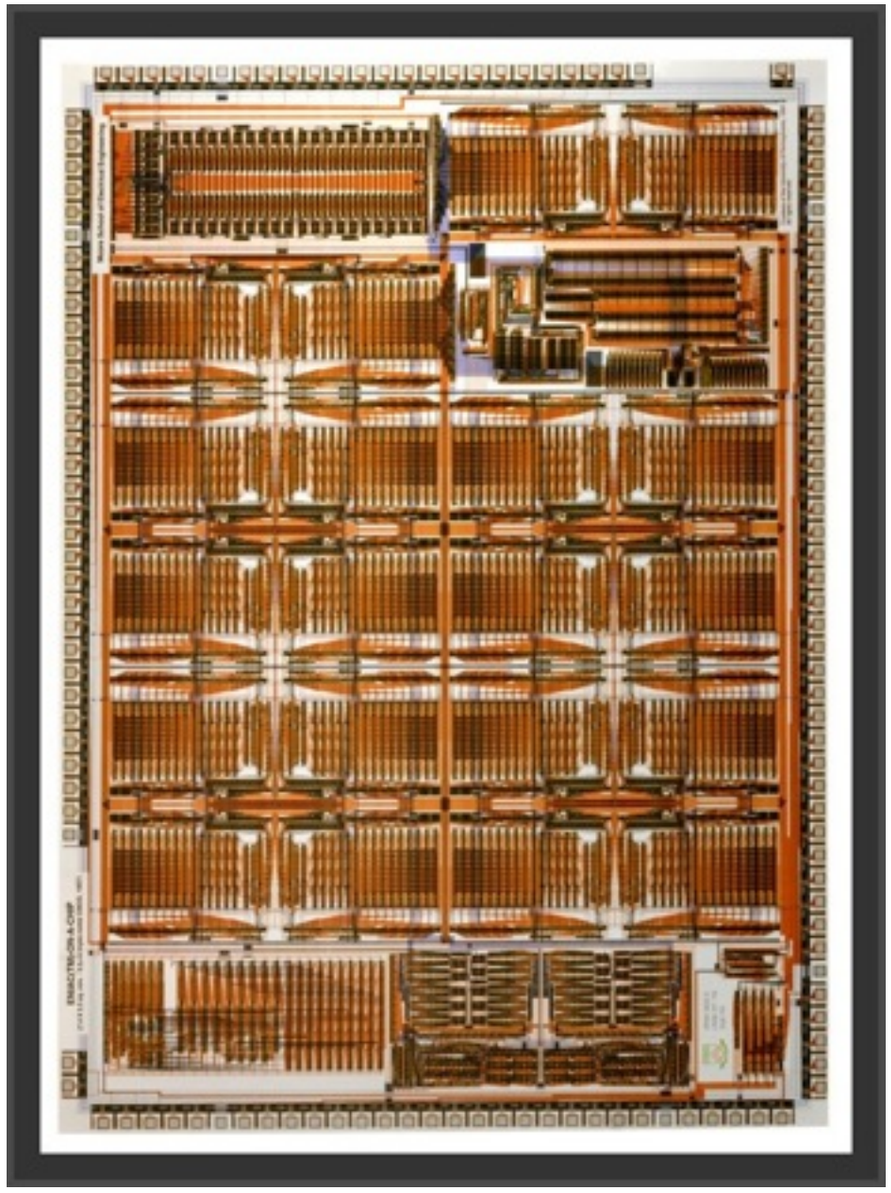
Other Lab Visits Planned

- ***Lab visit to see multicore server***
- ***(have requested) Downtown Data Center Tour***

So let's see some multicores (pics)
ENIAC vs. Single-chip ENIAC



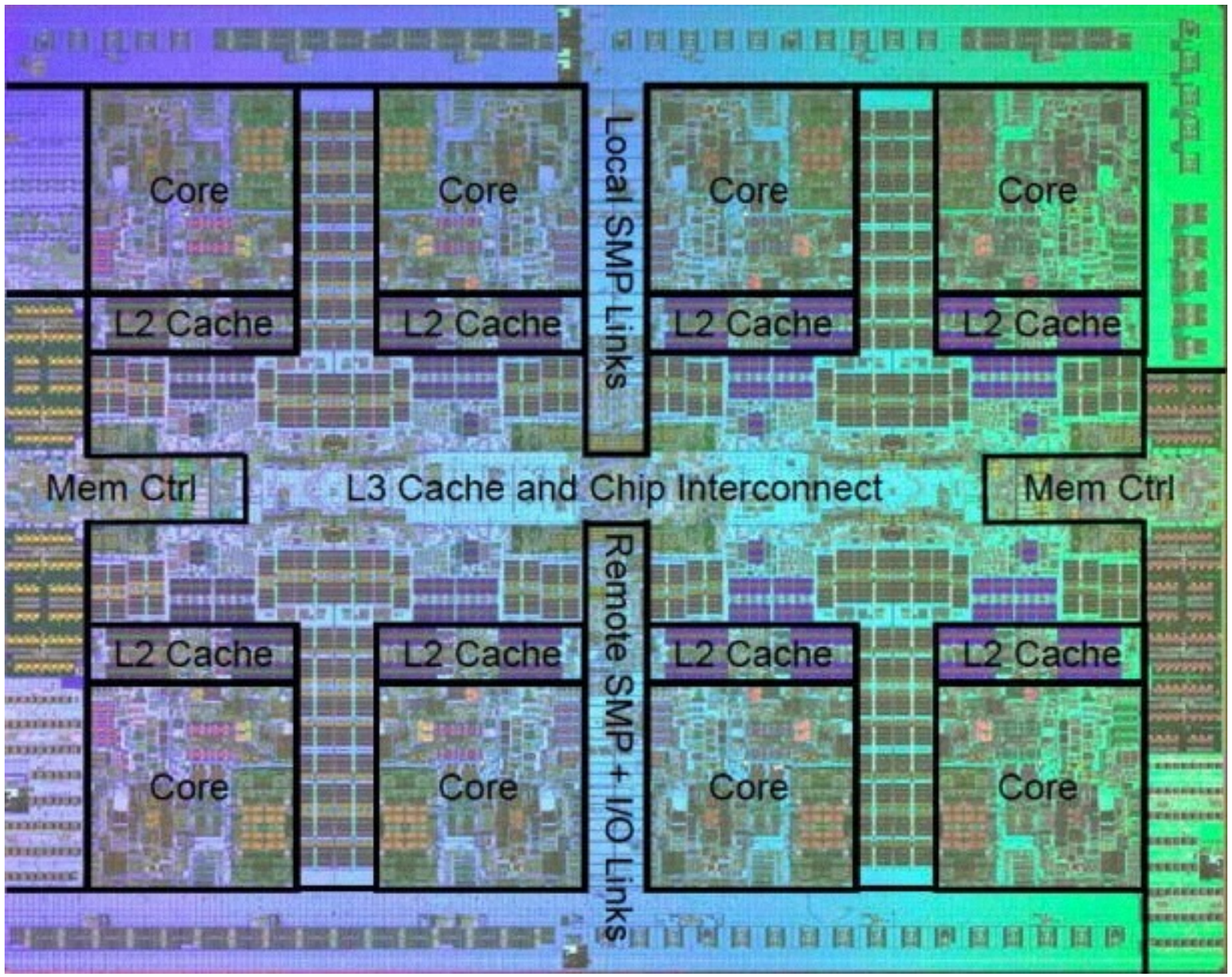
http://www.google.com/imgres?imgurl=http://upload.wikimedia.org/wikipedia/commons/4/4e/Eniac.jpg&imgrefurl=http://en.wikipedia.org/wiki/ENIAC&h=196&w=257&tbnid=HVRmUhC7IBNKrM:&zoom=1&tbnh=153&tbnw=200&usg=__0WcA-0E2Ml5Kf7hvy8PKq7Abk7M=&docid=jpc3gMQF2wmnGM&itg=1&sa=X&ei=alPeU-G6G4WUyASwg4HIBw&ved=0CJwBEPwdMAo



http://diephotos.blogspot.com/

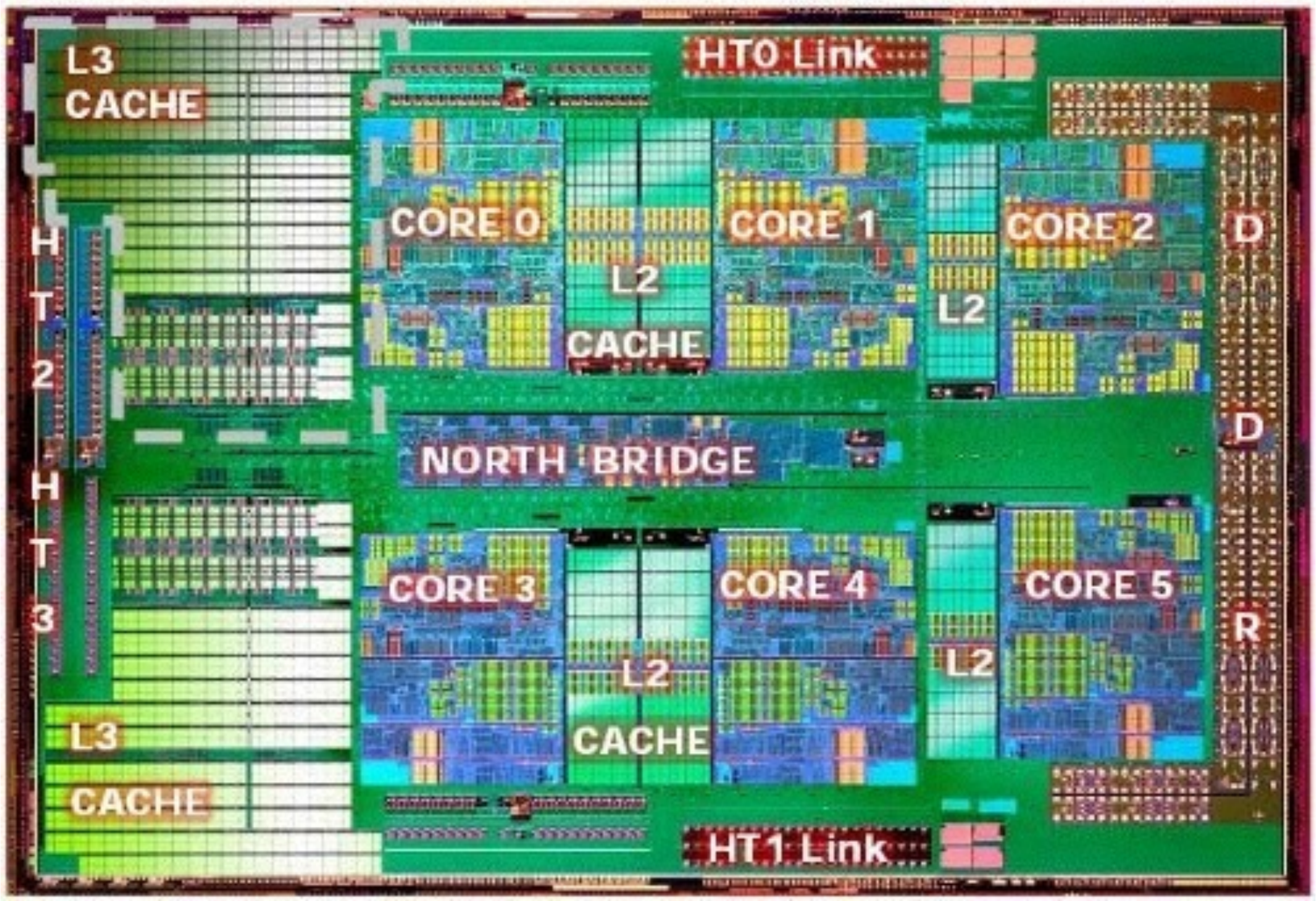
ENIAC-ON-A-CHIP PHOTO COURTESY OF PROFESSOR J. VAN DER SPIEGEL, UNIVERSITY OF PENNSYLVANIA (C) 1997

So let's see some multicores (pics): 8-core Power-7 showing chip interconnect, Mem Controller, L3, and L2, and I/O Links



https://www.google.com/search?q=ibm+power+8+die+photo&espv=2&tbm=isch&imgil=MfH0d1yPgVUZXM%253A%253BdBAQYM57NmOdaM%253Bhttp%25253A%25252F%25252Farstechnica.com%25252Fgadgets%25252F2009%25252F09%25252Fibms-8-core-power7-twice-the-muscle-half-the-transistors%25252F&source=iu&usg=__NswhZ68jjewPgSJkP2Uskn-LUVM%3D&sa=X&ei=5FPeU9jQF5enyATU9oDADw&ved=0CDkQ9QEwBg&biw=1440&bih=713#facrc=_&imgdii=MfH0d1yPgVUZXM%3A%3BRcT0GpPCIS4CMM%3BMfH0d1yPgVUZXM%3A&imgrc=MfH0d1yPgVUZXM%253A%3BdBAQYM57NmOdaM%3Bhttp%253A%252F%252Fcdn.arstechnica.net%252Fpower7_ars.jpg%3Bhttp%253A%252F%252Farstechnica.com%252Fgadgets%252F2009%252F09%252Fibms-8-core-power7-twice-the-muscle-half-the-transistors%252F%3B610%3B483

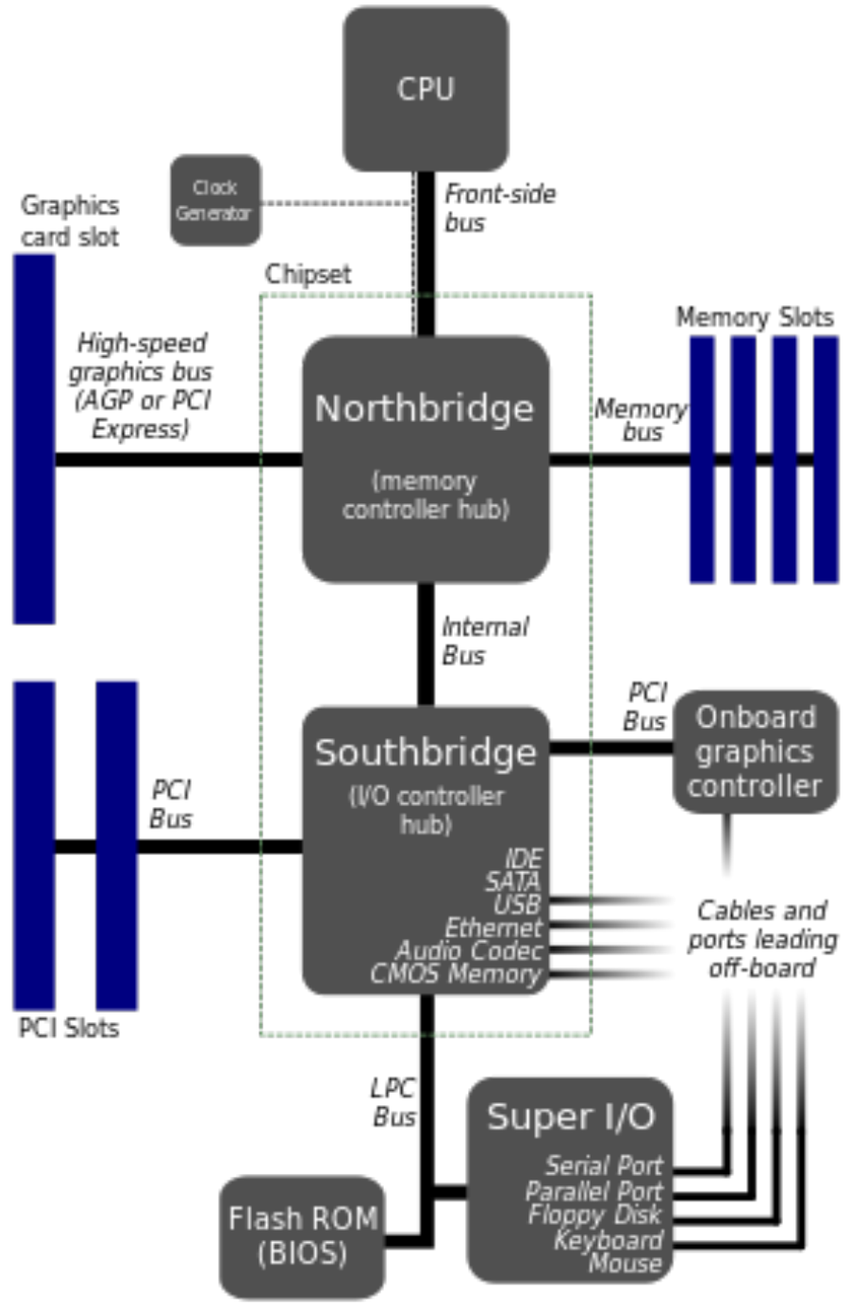
So let's see some multicores (pics)
AMD Magny Cours (6-core, showing the four Hypertransport Links HT0-HT3, North Bridge, L3,L2)



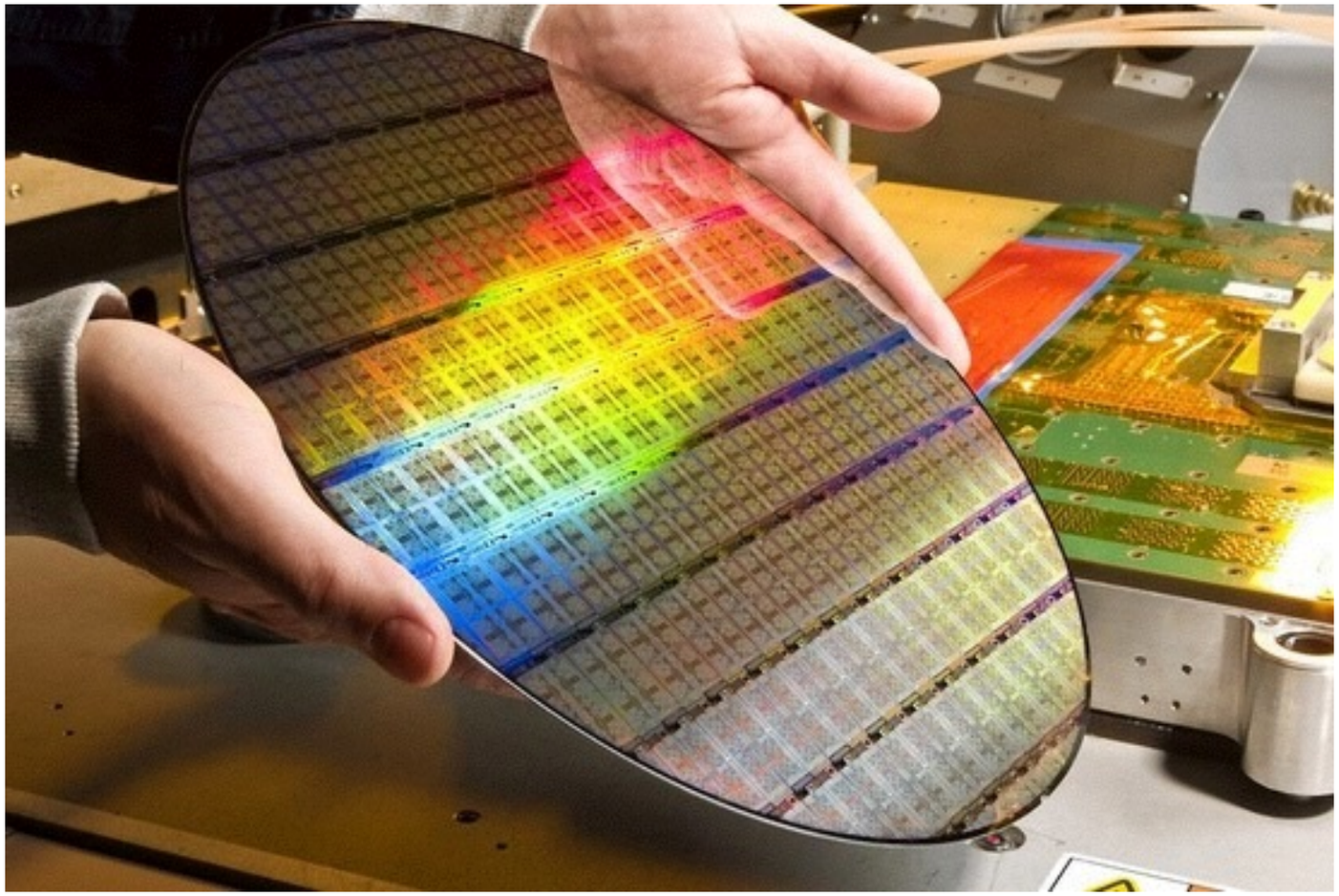
https://www.google.com/search?q=ibm+power+8+die+photo&espv=2&tbm=isch&imgil=MfH0d1yPgVUZXM%253A%253BdBAQYM57NmOdaM%253Bhttp%25253A%25252F%25252Farstechnica.com%25252Fgadgets%25252F2009%25252F09%25252Fibms-8-core-power7-twice-the-muscle-half-the-transistors%25252F&source=iu&usg=__NswhZ68jjewPgSJkP2Uskn-LUVM%3D&sa=X&ei=5FPeU9jQF5enyATU9oDADw&ved=0CDkQ9QEwBg&biw=1440&bih=713#facrc=_&imgdii=MfH0d1yPgVUZXM%3A%3BRcT0GpPCIS4CMM%3BMfH0d1yPgVUZXM%3A&imgrc=MfH0d1yPgVUZXM%253A%3BdBAQYM57NmOdaM%3Bhttp%253A%252F%252Fcdn.arstechnica.net%252Fpower7_ars.jpg%3Bhttp%253A%252F%252Farstechnica.com%252Fgadgets%252F2009%252F09%252Fibms-8-core-power7-twice-the-muscle-half-the-transistors%252F%3B610%3B483

Some of the subsystems are explained here

- ***http://en.wikipedia.org/wiki/Northbridge_%28computing%29***



So let's see some multicores (pics)
AMD Magny Cours (6-core, showing the four Hypertransport Links HT0-HT3, North Bridge, L3,L2)



A Si wafer with power chips on it.
(What do we do with the chips that fall beyond the perimeter ?!)

https://www.google.com/search?q=ibm+power+8+die+photo&espv=2&tbm=isch&imgil=MfH0d1yPgVUZXM%253A%253BdBAQYM57NmOdaM%253Bhttp%25253A%25252F%25252Farstechnica.com%25252Fgadgets%25252F2009%25252F09%25252Fibms-8-core-power7-twice-the-muscle-half-the-transistors%25252F&source=iu&usg=__NswhZ68jjewPgSJkP2Uskn-LUVM%3D&sa=X&ei=5FPeU9jQF5enyATU9oDADw&ved=0CDkQ9QEwBg&biw=1440&bih=713#facrc=_&imgdii=_&imgrc=eexyFksi-La5bM%253A%3BVQHbqk6IdRK0MM%3Bhttp%253A%252F%252Fsi.wsj.net%252Fpublic%252Fresources%252Fimages%252FMK-CF317_IBMGOO_G_20130805180127.jpg%3Bhttp%253A%252F%252Fonline.wsj.com%252Fnews%252Farticles%252FSB10001424127887323420604578650412719931232%3B553%3B369

Nasty Realities: Cracks, Aging, Metal Migration, Variability, Negative Temperature Bias Instability, Alpha Particle Strikes, Inability to avoid gates/flops failing due to inadequate margins, Dynamic Voltage and Frequency Stepping Noise, …

- *Cracks develop in metal wires, vias (connections thru holes)*
- *Chips age in many ways, including*
  - *Thermal stresses that accelerate cracks*
  - *Then the current density goes up*
  - *That leads to more thermal stresses, etc*
  - *Distributing workload on a multicore to even out the temperature profile is a good idea*
  - *NBTI causes the transistor threshold voltages to change*
    - *This causes the voltage needed to turn off/on a chip to change*
  - *Dynamically lowering frequency and voltage often saves energy*
    - *Lowering voltages : margin lower*
    - *Lowering frequency or even turning off clock (clock gating) : possible and often done*
- *According to Rob Aitken (ARM) and others present at a recent NSF/SRC workshop*
  - *Binning is done to recover and resell partially working multicores*
    - *E.g. IBM Cell Broadband Engine used to be sold as 5-core chips*
      - *The yield was so low that unless this was done, finding perfect 6-core parts was difficult (can't throw away partially working chips..)*
- *Imagine drones and other energy-critical systems that fly high (so are more prone to ambient noises and radiation also) : their computations cannot go astray*
- *The same for computer security applications - e.g. securing cars etc*
- *Future cars will be Chok-Full of Electronics ; Despite there being power, parasitic power harvesting can reduce required wiring*

Power and Energy

- *Power is rate of energy burn / production etc.*
- *Heat is proportional to power (not energy) because if you produce too much energy too quickly (i.e. high power), there isn't time for that heat energy to dissipate —> temperature rises*
- *Voltage times Current = Power*
  - *10 A at 1 V and 1 A at 10 V —> same power*
    - *But you need thicker wires to ship the former power*
  - *This explains why we step up voltage to 1000 KV and such before shipping long distance*
    - *(Low current, so wire of same size has less I^2 R loss)*
  - *Why HV DC is preferred : there isn't that much capacitive charge / discharge loss*
  - *Power Plant in Delta Utah ships HVDC for illuminating LA / NV ..*
  - *They get the electricity - we burn coal*
  - *Back pressure : So Delta is changing to Natural Gas (else CA can't buy because of their cleaner energy mandate)*

Power Line at Delta Utah

It is a bipolar overhead power line 488 miles (785 km) long, and can transfer a maximum power of 2,400 megawatts at ±500 kV. The part of the line that travels through the Mojave Desert is followed by many other AC 500 kV lines and Interstate 15.

2.4GW/500KV
=
4800 A

This is crazy!
(but w/o
high voltage,
current goes up
even more
and I^2 R
loss
will be huge)

The Intermountain HVDC power line. Picture taken in Utah along U.S. Route 50

Energy Consumption of Data Centers and HPC Installations

- ***See the TOP 500 list (as of Aug 3, 2014)***
  - **http://www.top500.org/lists/2014/06/**

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Super Computer Center in Guangzhou China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	**Piz Daint** - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9	2,325
7	Texas Advanced Computing Center/Univ. of Texas United States	**Stampede** - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462,462	5,168.1	8,520.1	4,510
8	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
9	DOE/NNSA/LLNL United States	**Vulcan** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
10	Government United States	Cray XC30, Intel Xeon E5-2697v2 12C 2.7GHz, Aries interconnect Cray Inc.	225,984	3,143.5	4,881.3	

We will now learn about microprocessor organizations, and also witness some of its manufacture (via Youtube videos)

- ***Here is how CPU chips are made***
  - **Courtesy Global Foundries**
  - **https://www.youtube.com/watch?v=qm67wbB5Gml**
- ***Here is how motherboards are made***
  - **Courtesy of Gigabyte Inc**
  - **http://www.youtube.com/watch?annotation_id=annotation_418185&feature=iv&src_vid=Va3Bfjn4inA&v=5vWrEmpRX_Q**
- ***Now read this excellent webpage (Wikipedia) on various parallel computing systems : http://en.wikipedia.org/wiki/Parallel_computing***
- ***This is a great set of notes on microprocessor organizations***
  - **http://www.cs.cmu.edu/~fp/courses/15213-s07/lectures/27-multicore.pdf**
  - **Notice that with Hyperthreading (a name coined by Intel), one thread can share FPU resources while another thread may be busy with Integer units**
- ***So finally, you can have cores with multiple threads running on it, but some of these threads can also run concurrently sharing resources on the chip***
- ***Read about Beowulf Clusters : http://en.wikipedia.org/wiki/Beowulf_cluster***

These are handy tools to see the performance of a Linux machine

- ***http://www.tecmint.com/command-line-tools-to-monitor-linux-performance/***

(Courtesy, Bailey et al. ; Also from Dongarra)
How the aggregate compute power of ALL the top 500 machines is surpassed by the 500th ranked machine a decade later.Will this trend continue? Is this a meaningful metric, given that "TOP 500" is evaluated by the ability to do LINPACK fast — not entirely representative of many real-world applications? Regardless, it is very interesting that so much progress used to happen over a decade.

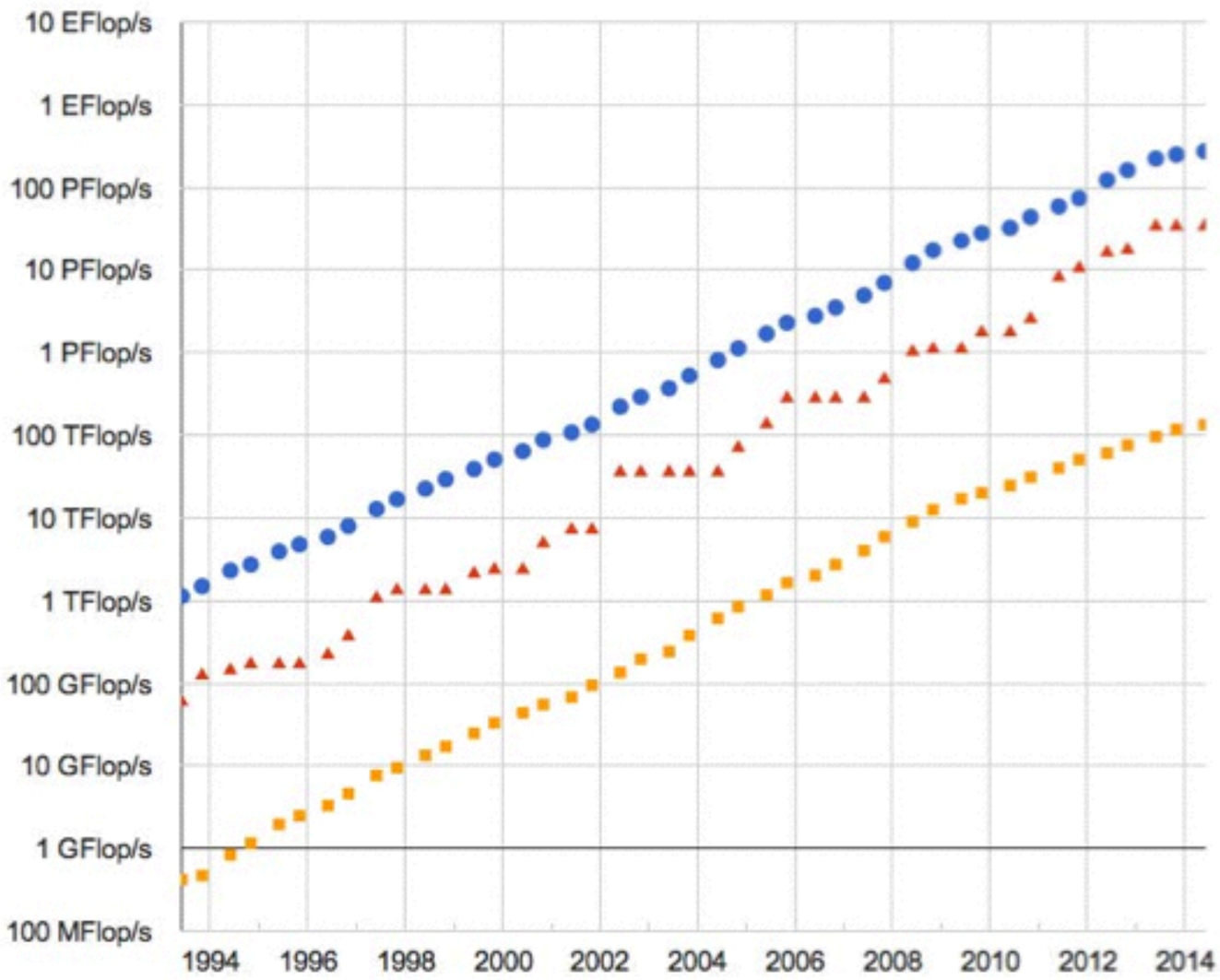


Figure 1: Performance of the Top 500 computers: Red = #1 system; orange = #500 system; blue = sum of #1 through #500.

thanks!

- ***www.cs.utah.edu/fv***

- ***Thanks to my former students who have taught me everything I know about FV and its relevance in the industry***