ECE 571 – Advanced Microprocessor-Based Design Lecture 22

Vince Weaver

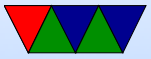
https://web.eece.maine.edu/~vweaver

vincent.weaver@maine.edu

28 October 2024

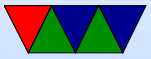
Announcements

- HW#7 was posted (short)
- Midterm Wednesday
- Interesting paper `http://ieeexplore.ieee.org/document/6757323/` the "Computing's Energy Problem" paper by Horowitz from ISSC2014.
- New ATX power supply proposals `https://www.theregister.com/2022/03/24/intel_updates_atx_psu_specs/`

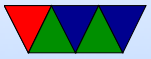


Midterm Notes

- One page of notes
- Benchmarking, skid
- Branch predicting. Know static, know 2-bit saturating counters
- Caches, there will be cache problems like homework (Sizing and also walking through example)
- Very brief / high-level questions on Virtual Memory, Processor Security, CPU power

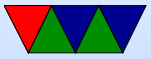


Security Wrapup from Last Time

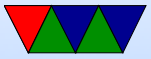


Other security things in modern architectures?

- Making RAPL counters less accurate, "energy filtering" to avoid info leakage
- SGX – Software Guard Extension
  - Encrypts memory on the fly by the CPU so no one except the owner of the code can see it (root, administrator, hypervisor all cannot see it)
  - Good for encryption keys and such, but also media companies like locking down video playing so you can't

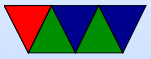


make copies (Ultra-HD Blu-ray depends on SGX)
  - Intel deprecated it on future chips starting in 2021
  - Side channels like described before can leak the state, lots of different attacks
- Signed firmware
  - Try to provide secure environment where only applications, operating systems, drivers, boot firmware, etc, all signed by central authority allowed to run

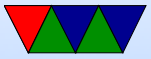


CPU Packaging Technologies

- Was reading this article: `https://www.theregister.com/2024/10/24/intel_amd_packaging/`
- Give brief rundown on how modern chips are made
- Billion dollar fabs, pure silicon, sliced into wafers, complex steps with high-freq light (EUV or x-rays these days) and poisonous chemicals
- Dies sliced up, defects reduce yield
- So many transistors too big for reticule
- AMD first and now Intel using chiplets, inside package

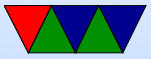


smaller chunks of chip connected. Downsides taking signals off chip, but flexibility in number of cores and cache

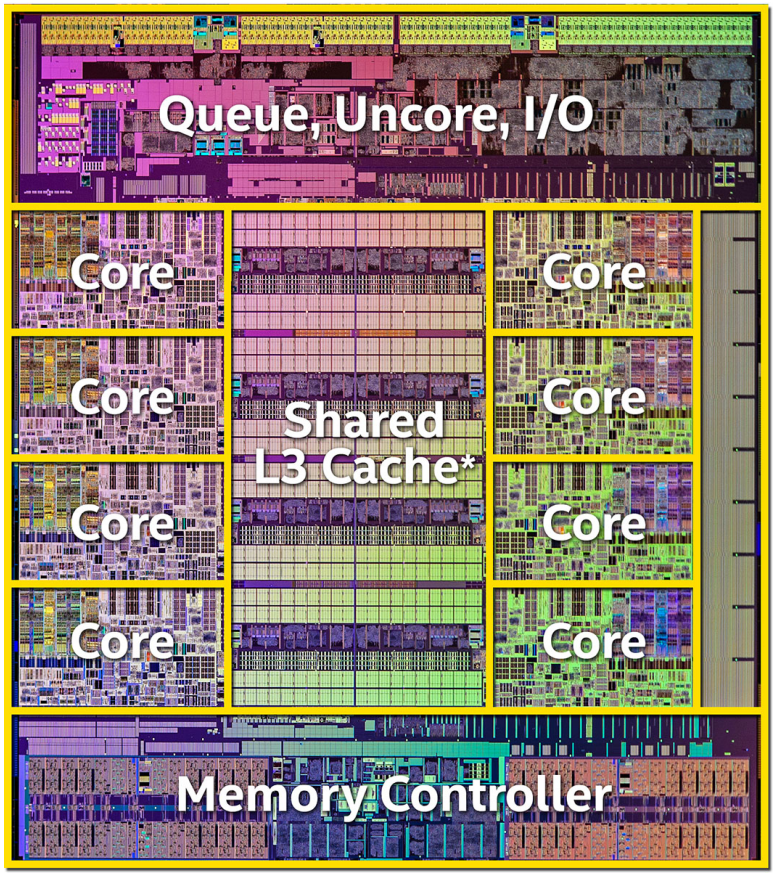


CPU Power and Energy

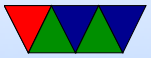
- Became a trendy thing to research in 1999-2002 timeframe.
- Before that usually concern was with performance.
- These days energy results are often reported as a core part of any architectural proposal, not as a separate issue.
- The results discussed here are academic and may or may not be implemented in actual chips.



Intel Haswell-E Die shot



Note which structures are big, using static power.

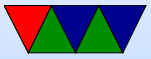


CPU Power Breakdown

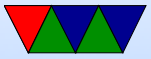
From Fan, Tang, Huan, Gao (ISLPED'05), Chinese Godson MIPS CPU

They gave numbers, but unclear of workload, if static or dynamic, etc.

- Cache 36%
- TLB 13%
- FALU 10%
- ROQueue 7%
- FMUL 6%



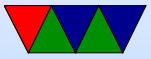
- Float reg 5%
- Gen reg 5%
- MUL 2%
- MCUControl 2%
- ALU 1%
- Other 13%



Thermal Concerns Too

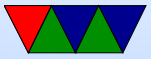
Power density exceed hot plate, approaching rocket nozzle

TODO: Find the Intel cite for this statement.



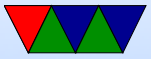
Methodologies Used in These Papers

It varies, but many of these are from simulations (sometimes validated). Anything from SPICE to "cycle-accurate" simulators.

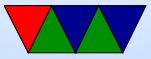


Clock Generation

- Driving high-frequency load against capacitance, trying to keep whole chip in sync.
- Extreme Case: Alpha 21264 H-tree, 32% of power?
- Half-frequency clocks (on both edge, so clock run half as fast) (Mudge 2001)
- Asynchronous – chips without clocks? Discuss Manohar's work, main problem lack of tooling/experience

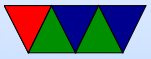


- Locally Asynchronous (Divide to multiple clock domains)



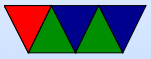
DVFS and other CPU Power/Energy Saving Methods

- A lot of related work
- Will focus on actual implementations rather than academic papers this time

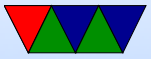


DVFS

- Voltage planes – on CMP might share voltage planes so have to scale multiple processors at a time
- DC to DC converter, programmable.
- Phase-Locked Loops. Orders of ms to change. Multiplier of some crystal frequency.
- Senger et al ISCAS 2006 lists some alternatives. Two phase locked loops? High frequency loop and have programmable divider?

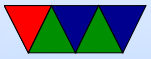


- Often takes time, on order of milliseconds, to switch frequency. Switching voltage can be done with less hassle.



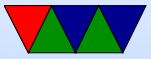
Compexity of DVFS

- Does Hardware or Software control?
- Turbo-boost
- Things like intel undervolt MSRs
- Things like AMD Ryzen CPCC2 which says which core is best for single-thread jobs



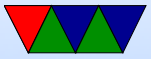
Adaptive Body Biasing

- Related to but not always considered part of DVFS
- Control voltage applied to body
- Change the threshold voltage
- Reduces leakage but slows performance



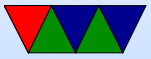
Cache Power and Energy

Large area, low-hanging fruit



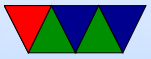
Decay Caches

- Kaxiras, Ho, Martinosi (ISCA 2001)
- Turn off cache lines not being used to reduce leakage
- DRAM cache with no refresh
- Decayed values can be re-fetched from memory. Tradeoff.



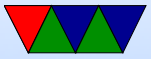
Drowsy Caches

- Flautner, Kim, Martin, Blaauw, Mudge. ISCA 2002.
- Move cold cache lines into "drowsy" mode.
  Lower power enough to hold state, not enough to lose contents. Reduce leakage. Better than decay as not lose data.
- Note: in Intel Volume 3b 17.17.5.2 it mentions certain C states might power down or otherwise turn off parts of cache.



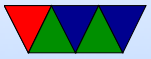
Adaptive Caches

- Albonesi (Micro 1999). Manually turn off ways in cache with an instruction.
- Size the caches



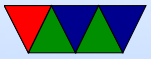
Cache Compression

- Dynamic zero compression for cache energy reduction (L Villa, M Zhang, K Asanović. Micro 2001).
- Cache Compression ("sign compression" – top bits) Energy savings 20% (simulated) (Kim, Austin, Mudge WMPI 2002)



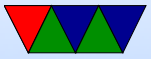
Banking and Filtering

- Filter cache, banking (only have half of cache active) (Mudge 2001)
- Slowing Down Cache Hits, Banked Data Cache. (Huang, Renau, Yoo, and Torrellas. Micro 2000.)
- Vertical Banking, Horizontal Banking (Su and Despain, ISLPED 1995).



Code Scheduling

- Can Schedule code for lower power.
- Better cache rates lower power. performance/power can go hand in hand. (Kandemir, Vijaykrishnan, Irwin)



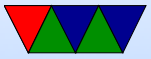
Branch Predictors

- Parikh, Skadron, Zhang, Barcella, Stan

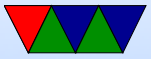
- 4 concerns:

  1. Accuracy. Not affect power, but performance
  2. Configuration (may affect power)
  3. Number of lookups
  4. Number of updates

- Tradeoff power vs time.

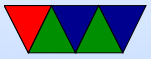


- brpred can be size of small cache, 10% of power
- Can use banking to mitigate



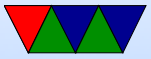
Branch Predictors

- can watch icache, not activate predictor if nobranches
- Pipeline gating, keep track of each predicted branch confidence. If confidence hits certain threshold, stop speculating. Show this may or may not be good.
- Integer code, large predictors good
- FP, tight loops, predictors not as important.



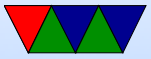
Branch Predictor Evaluation

- (Strasser, 1999). Simulation, small branch predictor can help energy.
- (Co, Weikle, Skadron) Formula for break even point. Leakage matters, what brpred hides is stall cycles.
- SEPAS: A Highly Accurate Energy-Efficient Branch Predictor (Baniasadi, Moshovos. ISLPED 2004). Once a branch prediction reaches steady state (unlikely to change) stop accessing/updating predictor, saving



energy.

- Low Power/Area Branch Prediction Using Complementary Branch Predictors (Sendag, Yi, Chuang, Lija. IPDPS 2008)
  Complementary Branch Predictor to handle the tough cases.

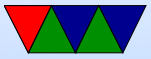


Prefetching

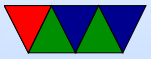
- Prefetching does not get looked at as closely.
  Various studies show it can be a win energy wise, but it is a close thing.

- (Guo, Chheda, Koren, Krishna, Moritz. PACS'04)
  HW Prefetch increase power 30%; have compiler help augment with hints, filters.

- (Tang, Liu, Gu, Liu, Gaudiot. Computer Architecture Letters, 2011).

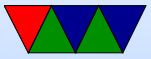


Mixed results.



TLB Energy

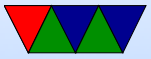
- Similar to caches
- Can you turn them off?
- Can you run w/o VM? uclinux?



TLB Optimization – Assume in Same Page

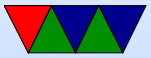
- Optimizing instruction TLB energy using software and hardware techniques (Kadayif, Sivasubramaniam, Kandemir, Kandiraju, Chen. TODAES 2005).
  Don't access TLB if not necessary. Compare to last access (assume stay in same page) Circuit improvements

- (Kadayif,Sivasubramaniam, Kandemir, Kandiraju, Chen. Micro 2002)
  Generating Physical Addresses Directly for Saving Instruction TLB Energy Cache page value.



TLB Optimization – Use Virtual Caches

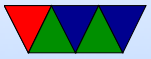
- (Ekman and Stenström, ISLPED 2002) Use virt address cache. Less TLB energy, more snoop energy. TLB keeps track of shared pages.



TLB Optimization – Reconfiguring

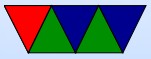
- (Basu, Hill, Swift. ISCA 2012) Reducing Memory Reference Energy with Opportunistic Virtual Caching
  Have the OS select if memory region physical or virtual cached.

- (Delaluz, Kandemir, Sivasubramaniam, Irwin, Vijaykrishnan. ICCD 2013) Reducing dTLB Energy Through Dynamic Resizing.
  Size TLB as needed, shutting off banks. Easier if fully-associative.



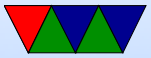
TLB Optimization – Memory Placement

- (Jeyapaul, Marathe, Shrivastava, VLSI'09) Try to keep as much in one page as possible via compiler.
- Energy Efficient D-TLB and Data Cache using Semantic-Aware Multilateral Partitioning (Lee, Ballapuram. ISLPED'03) Split memory regions by region (text/data/heap). Better TLB performance, better energy.



Bus Protocols

- Bus Protocols
- Cache-Coherence Protocols



Busses

- Grey Code, only one bit change when incrementing. Lower energy on busses? (Su and Despain, ISLPED 1995).
- PCIe
  - Linux-state power management
  - Shut down bus when idle
  - Problem is latency when waking up