# ECE 571 – Advanced Microprocessor-Based Design Lecture 20

Vince Weaver

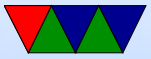
http://www.eece.maine.edu/~vweaver

vincent.weaver@maine.edu

4 April 2013

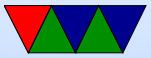
# Project/HW Reminder

- Reminder: Homework #4 due on Friday.
- Progress with Disabling Hardware Prefetch on Cortex A9
- Intel x86 machine available by Monday



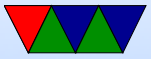
# TurboBoost

- Nehalem/Ivy Bridge/Sandy Bridge (AMD has similar Turbo CORE)
- Some Core2 had similar “Intel Dynamic Acceleration”
- Kicks in at highest ACPI Pstate
- “Dynamic Overclocking”



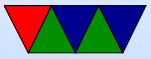
# TurboBoost – from HotChips 2011 Slides

- Monitors power, current, thermal limits, overclocks
- 100 uarch events, leakage function of temp and voltage
- P1: guaranteed stable state
  P0: turbo boost, maximum possible
- 12 temp sensors on each core
- PECI – an external microcontroller, used to control fans, package power

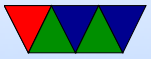


# TurboBoost example

- From WikiPedia Intel_Turbo_Boost article
- Core i7-920XM
- Normal freq 2.0GHz
- 2/2/8/9 – number of 133MHz steps above with 4/3/2/1 cores active
- 2.26GHz, 3.06GHz, 3.20GHz

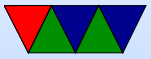


# Non-x86 Power Saving

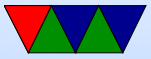


# IBM EnergyScale

- Thermal reporting
- Static and Dynamic Power Save
- “Power Folding” – reduce the number of CPUs reported to the OS until they are all busy
- Power Capping (like RAPL)
- Fan Control – Avoid “over-cooling”



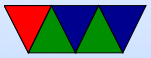
- Processor Nap – 2ms to wake up
- Processor Winkle (as in Rip Van) – 10-20ms to wake up, 95% of power



# ARM Cortex A9 (Pandaboard)

- Cortex-A9 Technical Reference Manual, Chapter 2.4 Power Management

- Energy Efficient Features
  - Accurate branch prediction (reduce number of incorrect fetch)
  - Physically addressed caches (reducing number of cache flushes)
  - Use of micro TLBs



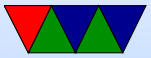
  - caches that use sequential access information? reduce accesses to tags
  - small instruction loops can operate without access icache

- Potentially separate power domains for CPU logic, MPE (multi-media NEON), and RAMs

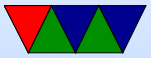
- Full-run mode

- Run with MPE disabled

- Run with MPE powered off



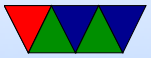
- Standby – entered with `wfi` instruction. Processor mostly shutdown except part waiting for interrupt
- Dormant – caches still powered
- Shutdown



# Pandaboard Power Stats

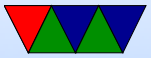
- Wattsuppro: 2.7W idle, seen up to 5W when busy
- http://ssvb.github.com/2012/04/10/cpuburn-arm-cortex-a9.html
- With Neon and CPU burn:

| | | |
|---|---|---|
| Idle system | 550 mA | 2.75W |
| cpuburn-neon | 1130 mA | 5.65W |
| cpuburn-1.4a (burnCortexA9.s) | 1180 mA | 5.90W |
| ssvb-cpuburn-a9.S | 1640 mA | 8.2W |



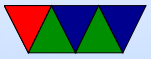
# Non-CPU power saving

- RAM
- GPU
- Ethernet / Wireless
- Disk
- PCI
- USB



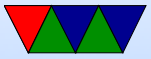
# DRAM

- Could teach a whole class on DRAM
- Tightly coupled to performance due to memory wall
- Commodity and churned out. Usually not interested in making changes to the underlying setup, usually just the interface or memory controller
- Memory controllers have migrated to the CPUs making that hard to change too

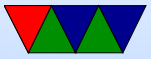


# DRAM – Mobile DRAM

- From Micron: "TN-46-12: Mobile DRAM Power-Saving Features", 2009
- Temperature-Compensated Self Refresh (TCSR) – Auto adjust refresh timings based on temperature
- Partial Array Self Refresh (PASR) – only refresh parts of RAM that have data in them
- Deep Power Down (DPD) – enable turning off the voltage generators when maintaining DRAM not needed



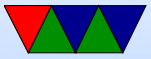
- Has equations for estimating power usage



# DRAM – Elsewhere

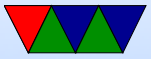
- Tom's Hardware. 2010. "How Much Power Does Low-Voltage DDR3 Memory Really Save?" Using low-voltage (1.25 or 1.35 rather than 1.5) DDR3 DRAM can reduce power by 0.5-1W. Slower performance settings, but not really noticeable.

- Linus Torvalds Rant from 2007: DRAM Energy not a prime concern. Just don't use FBDIMMs if you want low-power.



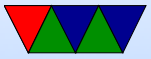
# DRAM – Recent Academic

- "Rethinking DRAM Power Modes for Energy Proportionality", Malladi et al, Micro 2012.
  - DRAM spends lots of time idle, but latency is so high for wakeup it cannot utilize powerdown modes
  - Reference 25% of data-center energy usage is DRAM?
  - Use LPDDR2 trades bandwidth for efficiency
  - Current modes involve turning off DLLs (Delay-locked loops?) which are slow to turn on again, 700ns+
  - some background on DRAM operation

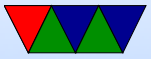


- Low-power mode sounds good, but then it takes 512 memory cycles of power to re-start (a lot of energy)
- Propose MemBLAZE. Moves clock generation out of DIMM and into memory controller, allowing fast wakeup

- "Towards Energy-Proportional Datacenter Memory with Mobile DRAM", Malladi et al, ISCA 2012.
  - Look at using LPDDR2 in servers rather than DDR3.
  - DDR3 often in "Active-idle" as many workloads do not allow sleep.

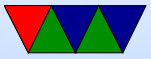


- "A Predictor-based Power-Saving Policy for DRAM Memories", Thomas et al, EuroMicro 2012.
  - Use a history based predictor to pick when to powerdown.
  - Say up to 20% of mobile devices and 25% of data center is DRAM
- "Rethinking DRAM Design and Organization for Energy-Constrained Multi-Cores", Udipi et al., ISCA 2010
  - DRAMs "overfetch" which hurts energy
- "A Comprehensive Approach to DRAM Power



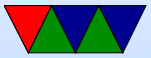
Management", Hur and Lin, HPCA2008.

- Throttling and Power Shifting – slowing down to fit in power budget
- Put DRAMs in low power mode – available commercially but no one seems to use this yet
- Simulate for Power5 and DDR2-533
- Modify the memory controller

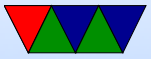


# GPU power saving

- From Intel lesswatts.org
- Framebuffer Compression
- Backlight Control
- Minimized Vertical Blank Interrupts
- Auto Display Brightness
- from LWN: http://lwn.net/Articles/318727/

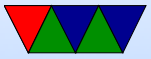


- Clock gating or reclocking
- Fewer memory accesses: compression.
  Simpler background image, lower power
- Moving mouse: 15W. Blinking cursor: 2W
- Powering off unneeded output port, 0.5W
- LVDS (low-voltage digital signaling) interface, lower refresh rate, 0.5W (start getting artifacts)



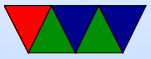
# Ethernet

- PHY (transmitter) can take several watts
- WOL can draw power when system is turned off
- Gigabit draw 2W-4W more than 100Megabit 10 Gigabit 10-20W more than 100Megabit
- Takes up to 2 seconds to re-negotiate speeds
- Green Ethernet IEEE 802.3az



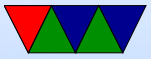
# WLAN

- power-save poll – go to sleep, have server queue up packets. latency
- Auto association – how aggressively it searches for access points
- RFKill switch
- Unnecessary Bluetooth



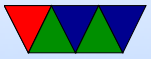
# Disks

- SATA Aggressive Link Power Management – shuts down when no I/O for a while, save up to 1.5W
- Filesystem atime
- Disk power management (spin down) (lifetime of drive)
- VM writeback – less power if queue up, but power failure potentially worse



# Soundcards

- Low-power mode



# USB

- autosuspend. Can sometimes cause issues
- off by default as some USB you disable don't come back