Smart Dust Wireless - 2004

* Power and energy related to volume
* 12pJ/instruction
* Independent subsystems, component level clock gating, processor halt mode with instantaneous power cycling, guarded ALU inputs, multiple buses
* Harvard
* No datapath pipelining
* Decoder pipelining using two delayed clocks; 1 cycle per instruction operation for all instructions
* Controller wakes up minimum hardware to execute a task
* 5 timers:
  + Two provide sample periods for sensor channels
  + Third invokes transmit block
  + Fourth invokes receive block
  + Fifth is a software timer that wakes up datapath
* Threshold comparison => SRAM with time stamp without invoking datapath

An ultra-low energy asynchronous processor for Wireless Sensor Networks - 2006

* 14pJ/instruction 170 MIPS at 1.2V and 2.7pJ/instruction 48 MIPS at 0.54V
* Reducing both power and energy requirements per computation activity (e.g. sensing operation, packet reception or transmission, routing action) permits to reduce the size of the battery and increase its lifetime, thus impacting both unit and deployment costs.
* 8 bit AVR
* Frequency and voltage scaling
* Zero wake up time (no PLL overhead)
* Voltage supply close to process threshold
* Flatter power spectrum and smaller voltage drops due to reduced power consumption peaks in the vicinity of clock edges
* Delays characteristics of different gates operating at same voltage vary in the same manner
* Clock tree replaced with async controllers
* Desynchronized netlist which outputs an asynchronous clock which can be used to drive synchronous peripherals
* Focus now is to reduce quantity of raw sensing data transmitted by nodes by increasing computational ability of each node and so power needs to be reduced. This reduces data traffic
* Desynchronization
  + FF to M-S
  + Matched delays for CL
  + Local Controllers
* Dynamic voltage and frequency scaling
  + Based on sampling the output of a signal which is forced to make a transition very close to the clock cycle, and slow down the clock frequency or increase the voltage supply if this critical sampling happens at the current voltage and frequency conditions.
  + Razor CPU
  + PowerWise
  + Delay line output is used to generate clock period
  + EMI emission reduction
* At 0.5 V the desynchronized circuit is 5 times more energy efficient.

An Ultra-Low Power Processor for Wireless Networks (SNAP/LE) - 2004

* Asynchronous processor
* Instruction set optimized for WSN applications
* 10pJ/instruction
* Hardware event queue and event coprocessors, which allow the processor to avoid the overhead of operating system software (such as task schedulers and external interrupt servicing), while still providing a straightforward programming interface to the designer.
* Low over-head transitions between active and idle periods.
* Circuits that don’t perform an operation have no switching activity
* Quasi delay insensitive (QDI) circuits. Isochronic fork assumption
* Characteristics:
  + Lower power sleep mode
  + Low over head wake up
  + Low power consumption while awake
  + Simple programming model
* Event driven
* No interrupts or exceptions. Only events.
* Time for event token to propogate through event queue is the awake time
* Fast and slow buses reduce capacitance

An Ultra Low Power System Architecture for Sensor Network Applications - 2005

* Event driven system
* Hardware Acceleration: Message Processor
* Regular events: Event processor, irregular events: microcontroller
* Non-pipelined
* Memory mapped interface for slaves
* Vdd gating
* Memory is partitioned so a block can be turned off if inactive
* Leakage and active power needs to be accounted for

Architectural and Circuit Design Techniques for Power Management of Ultra-Low-Power MCU Systems - 2013

* Reduce power in sleep and active modes and in transistion
* By exploiting the correlation between system clock speed and system power demand, this LDO digitally adapts its maximum current drive capability. To achieve fast and energy-efficient wake-up, this LDO does not need any external capacitance. It instead relies on the intrinsic capacitance of the MCU digital core amounting to overall 3nF.
* State-of-the-art MCU systems for this purpose make use of multiple LDOs, each optimized for a dedicated load condition, resulting in a slow and complex transition between system operating modes.
* Disable LDO during sleep
* A fully integrated LDO with low O/P capacitance
* Discrete load adaptive scheme which makes energy per cycle independent of clock speed

Hierarchical Power Distribution and Power Management Scheme for a Single Chip Mobile Processor

* Multi-chip organization
* 20, power domains, partial power-off
* Hierarchical power domains
* Reduced rush current

Memory Compression

Multi-Core for Mobile Phones

* Baseband processor and application processor
* Power consumed by Power Amplifier and RF
* 100GOPS from 1W
* Multiple small cores at a lower clock frequency is more power efficient
* Heterogeneous multi core
* Worload will surpass clock so need programmable cores

A 90nm Low-Power FPGA for Battery-Powered Applications

* Voltage Scaling
* Low Leakage for SRAM configurations (mid oxide thickness)
* Power gating
  + Gating per tile
  + NMOS
  + Ways to mitigate leakage paths
* 100ns standby mode

IA Processor

* Mobile internet devices
* Power efficient algorithms
* Intel Deep Power Down Technology [3] which allows for a majority of the CPU functionality to be powered down except for an on-die array that holds the micro-architectural state with very fast entry/exit times (100 us)
* Different power states (C0, high power to C6, low power)
* Power gating, shutdown PLL’s, flush L1 cache.
* Average power is 220mW and Idle power is 80mW.
* In C-0 high frequency mode (HFM) and low frequency mode (LFM), the processor can operate at its maximum frequency and its minimum frequency, respectively. In C-1 power state the core clock is power-gated and the L1 caches flushed resulting in lower dynamic power; exit latency is under 1 s. In C-4 power state, the 2 PLLs are shut down, the L1 caches are flushed leading to further dynamic power reduction; exit latencies are in the order of 30 s. Finally, in C-6 power state, the state of the machine is stored in an on-die SRAM (built as a 1-read, 1 write ported register file) and the core power supply is shut down resulting in the lowest power; exit latencies are in the order of 100 s.
* Gridless topology that routes clocks to places if required
* Register file design for all core arrays. Fine granularity sleep on word-line drivers with double stacked PFETs with negligible wakeup times.
* The sleep signal is generally shared across 8 entries. When any one of the 8 entries is going to be accessed, the sleep is de-asserted. When none of the entries in a group-of-8 is being accessed, the sleep is asserted and turns off the sleep PMOS transistor resulting in lower leakage due to transistor stacking.
* When the ROMis idle, the “slp” signal is asserted by the control logic. This causes the secondary pre-charge (on the intermediate bit-line) to be turned off, thereby floating the intermediate bit-line thus reducing leakage substantially; wake-up time performance impact is very small and comprehended in the cycle time analysis.
* L2 cache: data arrays remain in sleep mode until a qualified HIT signal is received from the TAG. This generates the SRAMWAKE signal for the relevant data sub-arrays and causes the memory cell to reach up to full voltage from its retention level; as a result only 4 out of 32 sub-arrays are on for any given time.
* See page 6
* Splitting power supply and only keeping relevant pins active reduces idle power.
* Addressing leakage power, word-line driver gating and floating bit-lines have been used along with complete shut-down of unused L2 cache sub-arrays.

Adaptive Body Bias

* Segmenting memory blocks to shutoff power when not in use. Supply voltage is scaled down in sleep state to provide reduction in power
* Periphery leakage power reduction is accomplished by separating the periphery power supply from array power supply. In this way, the core logic voltage is applied to the non-bit cell portion of the memory thus reducing power while the array power is controlled separately to ensure bit stability.
* SmartPriMer: generates power efficient synthesizable RTL and UPF compliant information.
* FBB to improve performance and RBB to reduce power. Only applied to parts of the chip

Energy Efficiency in a Mobile Processor that uses non-volatile memory

* Non-volatile memories reduce leakage current in standby mode.
* Resistive random access memory (RRAM) is considered to be one of the most promising emerging NVMs due to high speed, small area, and low power consumption.
* Retains state on power off but has slow write speed as compared to SRAM
* In RRAM, application of high voltage can set or reset a conduction path in a dielectric.
* Crosspoint structure increases memory density

Energy Consumption in mobile computing

* Coming together of hardware and software principles
* Identify components that use battery and scenarios as well (making a call)
* Screen consumes most power and GSM in suspended state
* Heterogenous multicore architectures goes with scheduling algorithms
* Software: Parallel Programming and scheduler policies
* Requests made by an application. For example, rendering pictures in different formats.
* Special domain names for webpages on mobile phones (reduces CSS and javascript)
* Parallel programming on several processors

Multicores: Use them or waste them

* Offlining allows the operating system to power-down individual cores, allowing the remaining cores to continue processing. DVFS, dynamic voltage and frequency scaling, provides for the reduction of CPU operating frequency (at the cost of performance), thus reducing dynamic power per the equation P α fV2.
* Choosing incorrect operating point (OP) is big. Choose frequency and number of cores.
* Lower frequency, more cores
* Better to have more cores running at a lower frequency and completing work faster than having offline cores
* Linux governor