Microprocessor power management

The Intel approach
The GHz Era
Increasing Power

Specint_rate2000; source: Intel; some data estimated.
March 12, 2003

Shift to PERFORMANCE per WATT in Notebooks

“The new processor delivers on Intel’s promise of faster performance and a battery-life boost measurable in hours.”

Bahtas

Intel inside

centrino

MOBILE TECHNOLOGY
Dual Core Shared Cache

2003
BANIAS
130nm

2004
DOTHAN
2x cache
90nm

2006
YONAH
Dual Core
Shared Cache
65nm

>2X INCREASE in Performance per Watt

Performance per Watt Heritage
Intel® Core™ Duo Processor
90 mm²
151M transistors

• Intel® Wide Dynamic Execution
• Intel® Advanced Digital Media Boost
• Intel® Advanced Smart Cache
• Intel® Smart Memory Access
• Intel® Intelligent Power Capability
• Intel® 64 Architecture
Delivered Performance = Frequency * Instructions Per Cycle (IPC)

Goal is higher performance and lower power

Power $\propto C_{\text{dynamic}} \times V \times V \times \text{Frequency}$

$C_{\text{dynamic}}$ is roughly a product of area and activity
“how many bits” * “how much do they toggle”
Delivered Performance = Frequency * Instructions Per Cycle (IPC)

Frequency is proportional to voltage, so frequency reduction coupled with voltage reduction results in cubic reduction in power.

Power $\propto C_{\text{dynamic}} * V * V * \text{Frequency}$
Processor Architecture 101

Delivered Performance = Frequency * Instructions Per Cycle (IPC)

Higher IPC usually results in wider data paths and/or more speculation: directly increasing C dynamic

Power $\propto C_{\text{dynamic}} \times V \times V \times \text{Frequency}$
Micro-op Reduction

Delivered Performance = Frequency * Instructions Per Cycle (IPC)

Fewer uops per instruction allows IPC to be increased while lowering C dynamic (less bits and less toggling)

Power = C_{dynamic} * V * V * Frequency
Techniques for Micro-op Reduction

- **ESP Tracker (Extended Stack Pointer)**
  - Execute Stack Pointer updates in dedicated hardware
  - *Intel® Core™ microarchitecture increases BW by 33%*

- **Micro-Op Micro-Fusion**
  - Single Uop representation of “multi-uop” instruction
  - *Intel® Core™ microarchitecture increase # instructions*

- **Macro-Fusion**
  - New technique in Intel® Core™ microarchitecture

*Techniques pioneered on Intel® Pentium® M processors*
Increasing Energy Efficiency

Specint_rate2000; source: Intel; some data estimated.
Intel® Intelligent Power Capability

- 65nm
- Strained Silicon
- Low-K Dielectric
- More Metal Layers

- Aggressive Clock Gating
- Enhanced Speed-Step

- Low VCC Arrays
- Blocks Controlled Via Sleep Transistors

- Low Leakage Transistors
- Sleep Transistors

Energy ADVANTAGE
- Mobile-Level Power Management
- Energy Efficient Performance

*Graphics not representative of actual die photo or relative size.* Intel Confidential.
Even during periods of high performance execution, many parts of the chip core can be shut off.

Example could be a SW memory initialization executing from front end with IQ operating as loop cache.
Intelligent Power Capability

Split Busses (core power feature)

By splitting buses to deal with varying data widths, we can gain the performance benefit of bus width while maintaining C dynamic closer to thinner buses.

Improved Energy Efficiency
Platformization of Power Management Architecture

- Integrating best features from Server and Mobile products
- Exposing more to the system

- **PSI-2**  Power Status Indicator (Mobile)
- **DTS**  Digital Thermal Sensors
- **PECI Interface**  Platform Environment Control
Power Status Indicator (Mobile)

- Processor communicates power consumption to external platform components
  - Optimization of voltage regulator efficiency
  - Load line and power delivery efficiency
Enabling Efficient Processor and Platform Thermal Control...

**DTS – Digital Thermal Sensor**

- Several thermal sensors are located within the Processor to cover all possible hot spots.
- Dedicated logic scans the thermal sensors and measures the maximum temperature on the die at any given time.
- Accurately reporting Processor temperature enables advanced thermal control schemes.
Platform Environment Control Interface (PECI)

Processor provides its temperature reading over a multi drop single wire bus allowing efficient platform thermal control.
Managing Power and Cooling Efficiency

Silicon:
- Moore’s law
- Strained silicon
- Transistor leakage control techniques
- Clock gating

Processor:
- Power scaling with load
- Policy based power allocation
- Multi-threaded cores

System Power Delivery:
- Fine grain power management
- Ultra fine grain power management

Facilities:
- Air cooling and liquid cooling options
- Vertical integration of cooling solutions

Power management from transistors to facilities
Increased Transistor Performance With 2\textsuperscript{nd} Generation Strained Silicon

\begin{itemize}
  \item ~16\% higher drive current plus ~20\% lower gate capacitance improves transistor switching speed by >20\% at same leakage level
\end{itemize}
Reduced Transistor Leakage With 2nd Generation Strained Silicon

Or, transistor leakage can be reduced ~5x at same drive current
Power Efficient Interconnect Includes 2nd Generation CDO

- **Low-k** carbon doped oxide (CDO) dielectric reduces interconnect capacitance (improved from 90 nm generation)

- Metal 8 layer is added for improved density and performance (1 more layer than 90 nm generation)

- Interconnect capacitance is reduced by use of low-k dielectric and by ~0.7x line length scaling

- Lower capacitance improves interconnect performance and reduces chip power
Sleep Transistors Reduce Leakage Power

SRAM Cache Sub-Block

NMOS Sleep Transistor

V_{DD}

V_{SS}

70 Mbit SRAM IR photos

Normal SRAM sub-block leakage

Sleep transistors shut off leakage in inactive sub-blocks

>3x SRAM leakage reduction with use of sleep transistors
Gate Dielectric Today is Only a Few Molecular Layers Thick

Individual Atoms

Polysilicon Gate Electrode

SiO$_2$ Gate Oxide

Silicon Substrate
High-k Demonstrated >100x Leakage Reduction

Benefits compared to current process technologies

<table>
<thead>
<tr>
<th></th>
<th>High-k vs. SiO$_2$</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacitance</td>
<td>60% greater</td>
<td>Much faster transistors</td>
</tr>
<tr>
<td>Gate dielectric leakage</td>
<td>&gt; 100x reduction</td>
<td>Far cooler</td>
</tr>
</tbody>
</table>
Leakage Control

Body Bias

Stack Effect

Equal Loading

Sleep Transistor

-Ve

-Vbn

Vdd

Vbp

+Ve

2-10X Reduction

5-10X Reduction

2-1000X Reduction
μArchitecture Techniques

- **Multi-threading**
  - Single Thread
    - Full HW Utilization
  - Multi-Threading
    - MT1: Wait for Mem
    - MT2: Wait
    - MT3: Wait

- Improved performance, no impact on thermals & power delivery

- Chip Multi-processing
  - Large Core
  - Cache
  - C1, C2, C3, C4

- Increase on-die Memory
  - Pentium® M
  - Pentium® II
  - Pentium®
  - 486

- Relative Performance vs. Die Area, Power
  - Multi Core
  - Single Core
Energy Efficient Performance

Performance Per Watt = System Performance

System Power Consumption

- Customers want to know how much power the system is consuming under real workloads – “Watts at the Wall”
- Use external power measurement tool
- System connects to Wall power outlet via Power Meter
CPU & electronics power in a platform is low
Must reduce power of other platform ingredients
Realization: Platform level

Power Delivery Improvements

**PSU**
- Typical PSU efficiency is 75%
- 80+% efficiency programs
  - EPA Energy Star (proposed), 80PLUS program, SSI specifications
- Technology exists for ~90% efficient PSUs
  - Cost remains a barrier to use, despite recuperation of cost within 1 year

Pathfinding for 90+% PSU
Light Load (5%) to Full Load (100%)
Fine-grain Power Management

- Improve response time
- Bring power supply closer

Power Supply Response

Usage
Time

Provide multiple supply voltages
Fine grain Vdd and Freq scaling

Fast
Slow

Slow

Fast

Cache

Core Hopping

For hot spot & power density management
Summary

- Energy Efficiency will continue to be a major design goal
- Holistic approach needed
  - Semiconductor process technology
  - Circuit design
  - Micro-architecture
  - Platform architecture
  - Power supply design
  - Software
Microprocessor power management

The ARM approach
What Consumers Care About

• Users want more features in their mobile devices:
  – MP3, Camera, Video, GPS...

• But also need long battery life
  – Convenient form factor, affordable price

• Battery technology is not evolving fast enough!
  – Need to manage power consumption
A typical low power SOC
A complex, asynchronous, mixed signal control system

Software

CPU

Power switches
isolation cells
Level Shifters
Save/restores

SOC Block

Power switches
Isolation cells
Level Shifters
Save/restores

VR

Button Press

Battery/Always On

PMU

Pin Straps

OTP

POR Logic

DMO
Range of Voltage-Control Techniques

- **Multi-Vdd (MV)**
- **MTCMOS power gating (shut down)**
- **Power gating with State Retention**
- **Dynamic or Adaptive Voltage Frequency Scaling (DVS, DVFS, AVS, AVFS)**
- **Variable V_{TH} (Back Bias – P/N)**
- **Low-VDD Standby**
Process Migration Alone Is No Longer Enough

The good old days

New reality

Everything improves significantly

Speed increases at the expense of energy consumption
RISC is a good starting point...

• Pure RISC can go too far in reducing the functionality of each instruction
  – More instruction fetches, less efficient cache usage (more external fetches)

• ARM retains some carefully chosen CISC like features
  – Conditional instruction execution
  – LDM/STM - Load/Store multiple registers
  – LDR/STR - Load/Store Register with base plus offset
  – Flexible “second operand” on ALU (Barrel Shifter)

• ARM7TDMI - 16 bit “Thumb” Instruction Set
  – High code density for system size/cost/power savings
  – Greater than 20% code size savings over 32-bit ARM code
  – Memory footprint comparable to 8/16-bit microcontrollers

The cheapest, fastest, most reliable components of a computer system are those that are not there!

--Gordon Bell
Low Power Is A System Problem

- Operating System with Software Policies
  - Managing the entry and exit to and from system sleep states
- System Level Control IP
  - Architectural design partitioning, hardware control
  - Sleep transition protocol management
- Library Level Support
  - Comprehensive low power components (ISO, Switch, Retention)
- EDA Software
  - Comprehensive automation yielding ultra low power design with optimal QoR
- Power Supply Management
  - External power supply control, power supply tolerances, etc.
- Process Technology
  - Trade-off between a high performance and low leakage process
Power Dissipation

\[ E = \int_{0}^{t} (V_{DD}I_{\text{leak}} + CV_{DD}^{2}f_{c})dt \]

Minimize \( I_{\text{leak}} \) by:
- Reducing operating voltage
- Fewer leaking transistors
- Reduce transistor leakage

Minimize \( I_{\text{switch}} \) by:
- Reducing operating voltage
- Less switching cap
- Less switching activity
Dynamic Power Optimization

• Dynamic Frequency Scaling (DFS)
  – Reduce operating frequency if possible
  – Reduces average power (but not task energy)
  – Eliminates NOPs

• Dynamic Voltage & Frequency Scaling (DVFS)
  – Requires DFS
  – Reduces voltage if frequency is reduced
  – Reduces task energy
  – Based on characterized frequency – voltage pairs (lookup table)

• Adaptive Voltage Scaling (AVS)
  – Closed loop optimization of VDD at run-time
  – Can save energy even at fixed frequency
ARM IEM* Principles

- Batteries have finite amounts of energy stored in them
- Running fast and then idling wastes energy

Only need to run just fast enough to meet the application deadlines
ARM IEM Technology

Hardware and software solution for energy management
Dynamic control of voltage and frequency scaling.

- IEM software connects to OS kernel and collects data.
- Multiple policies categorize the software workload.
- Prediction of future performance requirement is made.
- Suitable operating point (Voltage and Frequency) is set.
ARM IEM System Implementation
ARM IEM System Implementation
Trends In Power Dissipation

- Static power dissipation can no longer be ignored
  - It became significant at 90nm and dominant at 65nm
- Leakage currents are rising fast
  - Must be controlled by circuit design and optimization tools
Leakage Currents

- Transistors are not perfect switches – they always “leak”
  - Especially the high performance (low Vt) ones

  \[ I_{\text{SUB}}: \text{Sub-threshold Leakage} \]
  \[ I_{\text{GATE}}: \text{Gate Leakage} \]
  \[ I_{\text{GIDL}}: \text{Gate Induce Drain Leakage} \]
  \[ I_{\text{REV}}: \text{Reverse Bias Junction Leakage} \]

- Currently sub-threshold leakage dominates
  - Multi-threshold and Power Gating most effective
- However gate leakage is becoming significant
  - Can be mitigated by high K dielectric material
Coarse Grain Power Gating

• **Power switches shared by many cells**
  – Reduced area and performance impact

• **Ring Based Switch Topology**
  – Rings of switches encapsulate the power down block
  – Good for legacy IP, non-intrusive for optimized blocks
  – Preferable if no state retention used – no always-on mesh
  – Can add significant area cost to existing IP

• **Distributed Switch Topology**
  – Distributed switch cells in power down block – smaller
  – Can be implemented as a sparse array, in rows or columns
  – Seems to provide better QoR and control of IR drop
  – Better trickle charge management during power-up
Managing Voltage Drop

- **A reduced supply voltage impacts performance**
  - Adding switches to a power network will necessarily induce a voltage drop in that network in addition to the voltage drop due to power distribution across the mesh
  - Must minimize this additional switch induced voltage drop through considered architecture of the switch topology and switch size
  - Voltage drop across switch also causes standard cells to operate slightly reverse biased if well is tied to VDD

- **Distributed switching can help limit voltage drop**
  - Provides a finer control resolution on the switch size and placement
  - Increase the switch density and size in power hot spots
  - Can reuse the always-on power networks (for switch control) to power retention registers and additional always-on logic (save & restore signals)
State Retention Considerations

- Three possible approaches to state retention
  - Software based state save and restore (OS driven)
  - Hardware based state save and restore via scan structures (via AMBA)
  - Hardware based local state retention with retention registers

- Choice of retention scheme dependant on a number of factors:
  - Area overhead of retention registers and size of state space to be maintained
  - Performance impact of retention registers
  - Energy cost for save and restore when saving state externally
  - Real time cost for save and restore when saving state externally

<table>
<thead>
<tr>
<th>Approach</th>
<th>Standby Leakage</th>
<th>Area Overhead</th>
<th>S/R Energy Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Gating Only</td>
<td>Power Switches &amp; AO Logic</td>
<td>Power Switches &amp; AO Logic</td>
<td>Complete Reset Required</td>
</tr>
<tr>
<td>Power Gating with Software Based Retention</td>
<td>Power Switches &amp; AO Logic</td>
<td>Power Switches &amp; AO Logic</td>
<td>State Restore via Software</td>
</tr>
<tr>
<td>Power Gating with Scan Based Hibernation</td>
<td>Power Switches &amp; AO Logic</td>
<td>Power Switches &amp; AO Logic</td>
<td>State Restore via Scan Shift From Memory</td>
</tr>
<tr>
<td>Power Gating with Local State Retention</td>
<td>Power Switches, AO Logic, Retention Registers</td>
<td>Power Switches, AO Logic, Retention Registers</td>
<td>Minimal as state maintained locally</td>
</tr>
</tbody>
</table>
In Summary:

- Manage power in all modes in which a design operates
  - Dynamic power during device operation including active leakage
  - Static power dissipation during standby

- Maintain device performance while minimizing power consumption
  - Meet most aggressive performance targets while minimizing power
  - Aggressive power optimization when running at reduced performance levels
  - Minimize impact to performance by employing aggressive low power techniques
  - Employ a number of low power techniques in a single processor implementation

- Aggressive techniques for power management
  - Dynamic power minimized through OS directed performance scaling
  - Dynamic power minimized through use of Multi-Vt and Multi-L libraries
  - Standby power minimized through power gating with state retention
  - Additional standby power savings through use of threshold scaling (bias)

- Dynamic voltage and frequency scaling at the system level for power control
  - Hardware and software solution for energy management

- Utilize RISC architecture