Intel 45nm Core 2 Microarchitecture

Presented By: Alec Wiseman
Agenda

- Microarchitecture Overview
- Efficiency Enhancements
- Enhanced Dynamic Acceleration Technology
- Results
Microarchitecture Overview

- Transistor Innovations
- Cache
- Wide Dynamic Execution
- Macrofusion
- Speeds
- Efficiency Enhancements
- Enhanced Dynamic Acceleration Technology
- Results
Transistor Innovation

- High-k Metal Gate silicon technology
- Hafnium-based high-k gate dielectrics
  - significant reduction in electrical leakage
- Approximately twice the transistor density
- Approximately 30 percent reduction in transistor switching power
- More than 20 percent faster transistor switching speed or more than 5 times reduction in source-drain leakage power
- Greater than 10 times reduction in transistor gate oxide leakage
- Considered to be the “biggest change in transistor technology since the introduction of polysilicon gate MOS transistors in the late 1960s”
Cache Innovations

- Uses 24-way set associative “smart” cache
- Between 6 and 12 MB shared L2 cache
  - 50% more than previous family
- Advanced Smart Cache
  - multi-core optimized cache
  - each core can utilize up to 100 percent of available L2 cache
  - when a core requires minimal cache requirements, a more active core can take over an inactive core’s share
**Wide Dynamic Execution**

- 14 stage pipeline
- Wider execution core
  - each core can process up to four instructions simultaneously (up from three)
- Improved branch predictor algorithm
- Advanced Dynamic Execution engine
- Deeper instruction buffers
  - provides greater execution flexibility
Macrofusion

- common instruction pairs combined into a single internal instruction (micro-op)
  - example: compare followed by a conditional jump
  - takes place during the decoding stage
- further facilitated by an enhanced ALU
  - single cycle execution of combined instruction pairs
- extended to micro-op fusion
  - macro-ops broken down into smaller micro-ops
  - micro-ops fused and executed together
- reduces the number of micro-ops handled by out-of-order logic by more than 10 percent
Speeds

greater than 3 GHz frequency in some versions

front-side bus up to 1.6 GHz

- greatly improved over previously available 1.066 and 1.333 GHz
Efficiency Enhancements

- Microarchitecture Overview
- Efficiency Enhancements
  - Deep Power Down
  - Server Idle Power Improvements
  - Desktop Idle Power Improvements
- Enhanced Dynamic Acceleration Technology
- Results
Deep Power Down (DPD)

- **Idle Power State**

- In mobile platforms, consuming several watts of power while idle degrades battery life significantly

- **Processor Running States**

  - **C0** - running state; the only state in which the processor is executing instructions

  - **C1-2** - higher numbered states consume less power with higher exit latency

  - **C3** - processor Phased Locked Loop (PLL) is shut down; turns off all clocks in the chip

  - **C4** - voltage applied to the processor in C3 lowered to reduce leakage

  - Intermediate states (e.g. C1E) achieve lower leakage with Vcc reduction yet maintain cache coherent state with low exit latency
In 65nm Core family, C5 state was implemented - Enhanced Deeper Sleep state

- Vcc reduced further, below cache retention voltage
- only enough voltage to retain core state
- leakage per transistor is low, but still accumulates to significant levels due to number of transistors
In DPD (45nm Core 2 family), critical processor state is saved in dedicated SRAM on-chip

- powered by I/O power supply for the chip (VccP)
- core voltage reduced to very low level via Voltage Regulator Module (VRM)
- equivalent to the Vcc core being powered off; does not consume power
upon a break event:

- processor signals the VRM to ramp the Vcc back up
- re-locks PLLs
- turns clocks back on
- internal reset to clear states
- restores core state from SRAM
- opens up L2 cache

all steps completed in 150-200µs in processor hardware

- transparent to operating system and power-management software
**Server Idle Power Improvements (CC3)**

- Targeted at single-socket and multi-socket workstations and servers
- Memory access from a core in a multi-socket environment generates snoops to all other cores and sockets
  - Snoop activity accounts for 30 percent of active core power
  - Idle state cores had to be woken up to respond to the snoop
  - Goal is to avoid snoops into idle cores
- Prior to the 45nm Core 2 family, idle cores were put into a snoop-able state - Core C1
Idle cores now can be put into a non-snoop-able state - Core C3

- first-level caches are flushed to L2 cache
- prevents cross-core snoops

Additional latency to enter CC3 is insignificant (less than 1µs)

- CC3 can replace CC1 with no effect on software and operating systems

- In cases where additional latency can’t be tolerated, CC3 state can be exposed by the ACPI interface as C2
  - Operating system power management can choose CC1 or CC3

- Expected power consumption saving of about 10 percent
Desktop Idle Power Improvements

- Focused on reducing the overall energy usage
- Deeper Sleep state
  - prior to the Penryn family, processors relied on autohalt and Stop Grant states
  - Penryn family processors use Deeper Sleep state to lower idle power consumption
- The processor communicates to the platform through the:
  - Graphics Memory Controller Hub (GMCH)
  - I/O Controller Hub (ICH)
- The GMCH and the ICH can then power off portions of themselves required to service processor requests when the processor tells them it is in the Deeper Sleep state
Deeper Sleep state reduces voltage to the processor

- processor cannot execute instructions
- state is still retained

Power consumption about 50 percent lower than Autohalt and Stop Grant states

Communicates Deeper Sleep state to VRM

- VRM can then shut down all but one phase
- Active phase is boosted in addition to power saved

Exposed in ACPI tables as C3 state

Longer latency to memory traffic and interrupt response

- dude to having to ramp up voltage before responding to memory snoops and interrupts

BIOS configured exit latency tuned down to minimum level to compensate

- no noticeable adverse impact
Enhanced Dynamic Acceleration Technology

- Microarchitecture Overview
- Efficiency Enhancements
- Enhanced Dynamic Acceleration Technology
  - DAT Overview
  - Enhancements
  - EDAT Principles
- Results
Power-management feature first introduced in 65nm Core 2 mobile processors

- Increases single threaded performance by using power headroom unused by idle cores

- Total power consumption remains within the specified TDP

- 10 percent frequency boost to single-threaded applications on mobile platforms
Enhancements

- Hysteresis mechanism allows tolerance of short wake-up intervals of the idle core
  - EDAT frequency retained
  - Reduced the number of transitions in and out of EDAT
  - minimizes performance loss in high-interrupt-rate workload
- Extended DAT support to quad-core mobile processors
  - achieved in the same way as Deeper Sleep state
    - core idleness information is shared
    - each site locally resolves whether or not to allow the processor to run at EDAT frequencies
EDAT Principles

- EDAT Frequency is fixed
  - typically 267 or 333 MHz over max core frequency

- Running half of available cores at EDAT frequencies while the remaining half remain idle will not exceed TDP limitations

- EDAT frequency is requested via the SpeedStep interface

- EDAT transition will only happen when the appropriate number of cores are idle and the operating system requests the highest performance (P) state

- If the operating system requests the P-state, the processor will run at guaranteed frequencies if enough cores are not idle
Results

- Microarchitecture Overview
- Efficiency Enhancements
- Enhanced Dynamic Acceleration Technology
- Results
  - Deep Power Down Results
  - CC3 Results
  - Deeper Sleep state and EDAT Results
Deep Power Down Results

- Benchmarked using MobileMark 2005 (MM05)
- Average power consumption reduced by 27 to 44 percent
- Testing done on Intel Customer Reference Board platform
- Fresh build of Windows XP
- Power measured by sensing voltage and current during the MM05 run and averaging it
CC3 Results

- Intel Customer Reference Board platform with Seaburg chipset and ESB 2
- SPECpower_ssj2008 benchmark
- Compared with C1E only enabled
- Power consumption reduced by 0 percent at complete idle to 10-20 percent at medium loads
- Power savings decreased at maximum loads due to time in C1E or CC3 decreasing
Deeper Sleep State and EDAT Results

- Deeper Sleep state
  - Measurements done under operating system idle conditions
  - Idle power consumption decreased by 40-60 percent

- EDAT
  - SPECInt2000, SPECFP2000, and two-threaded counterparts used for benchmarks
  - Performance increased by 5 to 8 percent
  - 11 percent frequency boost over guaranteed frequency
  - Hysteresis measured with SPEC2000 workload with 1,000 interrupts/sec timer
    - Used to tune hysteresis post-silicon
    - Recovered over half of lost performance gains due to high interrupt rates
Questions?