



Intel[®] Technology Journal

The Original 45nm Intel Core™ Microarchitecture

Power Management Enhancements in the 45nm Intel[®] Core™ Microarchitecture

Power Management Enhancements in the 45nm Intel® Core™ Microarchitecture

Jose Allarey, Mobility Group, Intel Corporation
Varghese George, Mobility Group, Intel Corporation
Sanjeev Jahagirdar, Mobility Group, Intel Corporation

Index words: power management, CPU Sleep, C states, idle power, frequency

Citations for this paper: Jose Allarey, Varghese George, Sanjeev Jahagirdar “Original 45nm Intel® Core™2 Processor Performance” Intel Technology Journal. <http://www.intel.com/technology/itj/2008/v12i3/2-paper/1-abstract.htm> (October 2008).

ABSTRACT

Intel® processors based on the original 45nm Intel Core™ microarchitecture, originally referred to by the codename Penryn, improved the energy efficiency and performance per watt of the Intel Core microarchitecture. This paper discusses the new technologies introduced in the Penryn family of processors that enabled lower idle power and higher performance levels.

The Penryn family of processors builds on the power-management capabilities of the Intel processors based on 65nm Intel Core microarchitecture, originally referred to by the codename Merom, and takes them to the next level of idle power reduction and multi-core Enhanced Dynamic Acceleration Technology performance. The Penryn family of processors took very aggressive goals for idle power reduction in mobile, desktop, and server platforms. All features have been demonstrated to be fully functional on silicon and meet or beat the expectations of power reduction and increased performance. In fact, most of these features have already been enabled on the products that are currently shipping. Furthermore, the Penryn family of processors introduced new processor sleep states and Dynamic Acceleration mode on Intel Core 2 Quad processors for the first time in Intel to enable the acceptance of quad core as a mainstream product.

INTRODUCTION

The Penryn family of processors, implemented in a 45nm high-k metal gate silicon process technology, is designed to fit a wide range of power envelopes and market segments. Energy efficiency (energy consumed for doing a unit of work) is significantly improved due to

process power scaling and innovative architectural power-management features. Power scaling enables better performance and higher power efficiency in most workloads, and the power-management features primarily enable lower idle power that leads to an overall reduction of platform energy consumption. Lower idle power helps improve battery life in mobile platforms, allows platforms based on the Penryn family of processors to meet or exceed Energy Star and other regulatory requirements for idle power consumption in desktop PCs, and lowers electricity and cooling costs for servers. The details of the 45nm high-k metal gate process technology and its power benefits are discussed in the last issue of the Intel Technology Journal [1,2] and hence are not covered here in detail. In this paper we focus on the architectural innovations in the power-management domain of the Penryn family of processors.

The Penryn family of processors builds upon the power-management capabilities of the core microarchitecture. The Advanced Configuration and Power Interface (ACPI) specification [3] describes the processor sleep states and performance states in detail. When the processor is executing instructions, it is in the C0 active state. The C1, C2 states, etc. are successively lower-power processor sleep states in which no instructions are being executed. P states are processor performance states defined by the processor frequency. The Penryn family of processors supports the Core Microarchitecture C states, P states, and thermal monitoring functions. In addition, the Penryn family of processors introduces several new key features and extends some of the existing mobile platform capabilities to desktop systems. Two key features introduced in the Penryn family are as follows:

- Deep Power Down (DPD), a new idle power state.

- Enhanced Dynamic Acceleration Technology (EDAT), a feature to increase Single Threaded (ST) performance by using the power headroom of the idle core. A simpler version of this feature is also available in the later versions of the Intel® Core™2 Duo 65nm processor.
- Power Management features extended to other segments are as follows:
- The Deeper Sleep state is now available in desktop platforms and in quad-core products.
- A version of the Deep Sleep state (called CC3 or Core-C3) state is now available in server platforms.
- The Enhanced Dynamic Acceleration Technology (EDAT) and Deeper Sleep state technology were extended to the mobile Quad-Core Extreme Edition product.

CHALLENGES

The Penryn family of processors was developed as a common core for the mobile, desktop client, workstation, and DP MP server platforms. The processor Thermal Design Power (TDP) in these segments ranges from 25 watts to 130 watts. Each platform has some common and some independent goals in power and thermal management. The challenge to the core development team is to anticipate the various requirements and design the support for all these platforms into the common core. For example, generic improvements such as reducing peak power and idle power consumption are applicable to all platforms. On the other hand, reducing energy usage during run time is a primary concern for the mobile and server platforms more than the desktop platform. Additionally, the mobile platform has tighter constraints on the peak power dissipation due to its form factor. Server platforms demand higher peak performance and are also capable of supplying and dissipating the power to support the peak demand conditions. The desktop platform usage mode has a lot of idle-ON time and standby time. Hence this platform has a specific requirement to reduce the idle and standby energy usage so that the end user will have a more energy-efficient, always-on always-available experience.

IDLE POWER IMPROVEMENTS

DPD technology

In mobile systems, battery life is an extremely important consideration. This is driving the need for low “average power” consumption in mobile processors. Designing high-performance mobile processors that consume approximately 30–40 watts during

normal operation but have extremely low-power consumption during idle is challenging. The leakage in the millions of transistors in a processor design adds up to several watts if not several tens of watts. Consuming several watts of power while idle degrades the battery life significantly.

In the days of the Intel 486™ processor there was only one type of idle state—the Autohalt state. Most of the clocks to the processor were stopped in this state; active power was cut down significantly, and this brought the total processor power consumption down because leakage was not an issue. Since then, however, leakage has gotten worse with every new process generation, and more aggressive power-management states (C states) have been added to processors to combat this issue while idle.

A brief outline of the various C states follows as an introduction to the DPD technique.

The processor running state is called the “C0 state” in Advanced Configuration Power Interface (ACPI) [3] nomenclature. The processor is not executing any instructions in C states other than C0. A higher-numbered C state generally consumes lower power at the expense of higher exit latencies than a lower-numbered one.

In the C3 state, the processor Phased Locked Loop (PLL) is shut down to turn off all the clocks in the chip. This, however, does not lower the leakage, since voltage is not changed. In the C4 state, the voltage applied to the processor in C3 is lowered to reduce leakage. Here, the voltage is lowered only to the point where state can be retained in both the core and the caches. Intermediate states such as C1E, which achieve lower leakage with Vcc reduction yet maintain the advantages of a cache coherent state with low exit latency, have been implemented in recent processors.

The mobile product of the Merom processor family has added a state called Enhanced Deeper Sleep state (referred to as the C5 state), in which Vcc is reduced even further, that is, below the cache retention voltages. In C5, the Vcc to the core is just high enough that the processor core retains its state. At these voltage levels, leakage per transistor is low; however, given that the processors have millions of such leaky devices, it still adds up to a significant power loss.

In DPD the critical state of the processor is saved in a dedicated SRAM on-chip that is powered by the I/O power supply for the chip (VccP), and then the core voltage is reduced to a very low level via the Voltage Regulator Module (VRM). Figure 1 shows the processor-VRM connectivity. At this point, it is

equivalent to the Vcc core being powered off, i.e., not consuming any power. Upon a break event such as an interrupt, the processor signals the VRM to ramp the Vcc back up, relock its PLLs, and turn clocks on. It then does an internal RESET to clear the states, restore the state of the processor from the dedicated SRAM, and open up the L2 cache. It then continues program execution from where it left off in the execution stream. All of these steps are completed in 150–200 us, all in the processor hardware, thereby making it transparent to the operating system and the existing power-management software infrastructure.

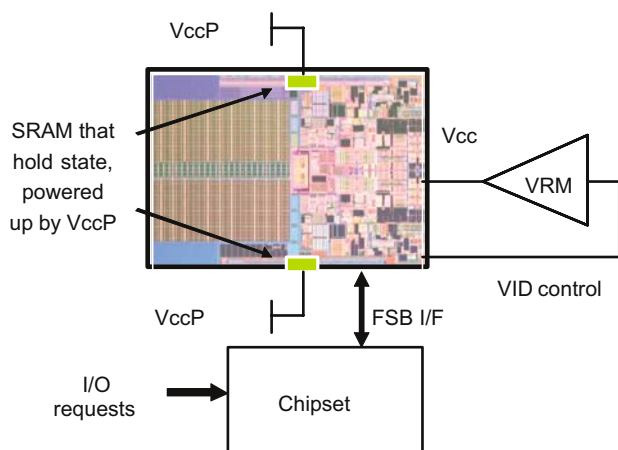


Figure 1: Voltage control for processor.

Because the entry and exit are managed entirely by the processor and do not require software assistance, DPD can be enabled under existing operating systems and platforms. DPD achieves a breakthrough for idle power since it is agnostic to Vcc min and provides a minimal power consumption state with a short exit latency. Preliminary results on silicon show that the DPD feature reduces processor average power by 27 percent to 44 percent as measured by the Mobile Mark 05 battery life benchmark.

Quad-Core idle power improvements

Until a few years ago, quad-core processors were used only in server platforms. Now they are available as mainstream desktop processors. In the Penryn family of processors, the Intel Core 2 Extreme processor offers four cores for the mobile platforms. To provide users with a battery life of greater than 2.5 h for DVD playback, it was necessary to reduce the idle power of the quad-core processor. Previous-generation quad-core processors supported only the Autohalt state and hence were not optimal for the mobile market segment.

Quad-core processors are implemented in a multi-chip package (MCP) configuration. The two dual-core

processor dice, referred to as the “Master” and “Secondary” sites, have independent PLLs, so the dice can run at different frequencies but share a common voltage plane. The Master die controls the voltage sent to the VRM and the voltage supplied to both dice. The secondary site coordinates its voltage requirements with the master die. The Penryn family of processors expands this coordination functionality for Deeper Sleep state support by using the same interface to communicate information regarding the sleep states of the cores. Each core pair on a die will resolve the correct idle state to enter. They will wait for the core pair on the other die to be ready to enter an idle state. At that point the Master die determines the resolved idle state and puts the package and the platform into that state.

Server idle power improvements (CC3)

The server versions of the Penryn family of processors are targeted for single-socket and multi-socket workstations and servers. In multi-socket platforms, a memory access from any core generates snoops to all other cores and sockets. Activity related to snoops of processor caches burns about 30 percent of active core power. If a core is in the idle state, it has to be woken up so that its caches can be queried for the data being requested by the snoop. After the snoop data has been returned to the requesting core, the woken core can return to the idle state. The wakeup and reentry negatively affects the idle residence and energy usage. By avoiding snoops into idle cores, this power can be saved.

In processors that predate the Penryn family of processors, the idle cores were put into Core C1 (CC1), which is a snoop-able state.

In the Penryn family of processors, idle cores can be put into Core C3 (CC3), which is a non-snoop-able state. The first-level caches are flushed into the L2 cache before putting cores into CC3. This prevents cross-core snoops and therefore the additional power burnt for snoops. Figure 2 shows how snoops are routed based on the cores’ CC states. The additional latency to enter CC3 is insignificant (less than 1 us). This allows CC3 to be used as a replacement for CC1 with no affect on software and operating systems. In cases in which the additional latency cannot be tolerated, the CC3 state can be exposed by the ACPI interface as C2 [3]. This lets the Operating System Power Management (OSPM) layer pick either CC1 or CC3, based on the latency and policy settings.

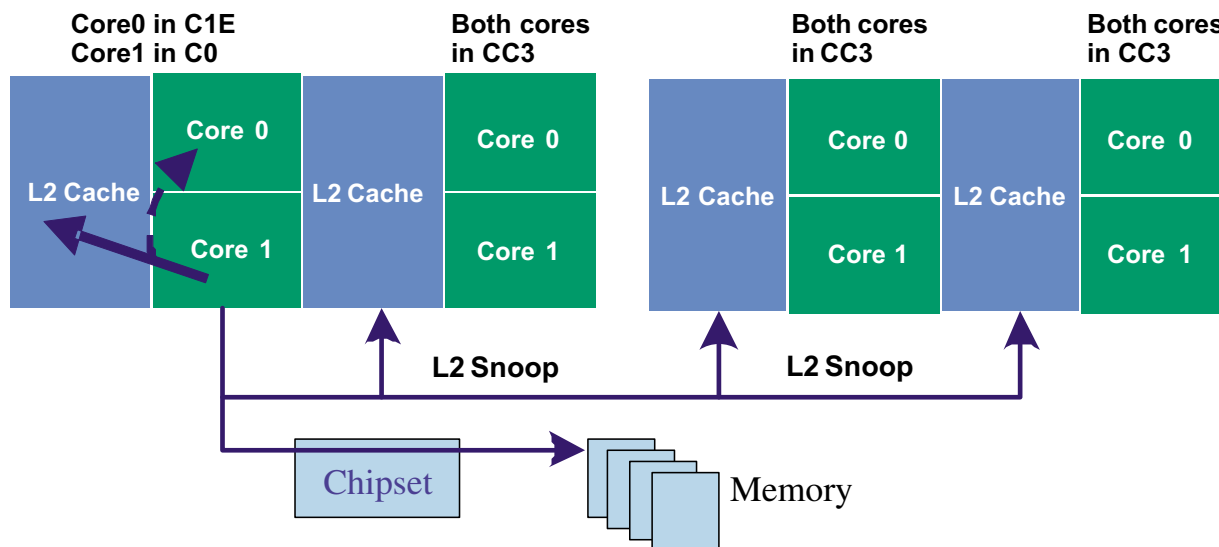


Figure 2: CC3 and cache snoops.

The estimated savings of CC3 relative to CC1 are directly proportional to the number of snoops that occur during the sleep state. In an idle scenario, the snoops in the system are very low. In a fully-loaded scenario, the residence in CC1/CC3 is very low. Hence, the savings in idle and fully-loaded scenarios is less than when the system is lightly loaded. On average, the expected power savings from this feature is about 10 percent.

Even though we have been discussing this as an idle power improvement, it is more than that. In today's typical server platform there are two, four, or more cores. Chances are one or more of these cores is in Idle, except for periods when the servers are under peak load conditions. The idle cores will enter CC3 and reduce the system power.

The other important aspect for server power savings is the fact that there are hundreds or thousands of servers in a server farm. A 10 percent reduction in energy usage not only translates into reduced energy costs but also increases the performance watt cubic feet ratio of the server farm.

Desktop idle power improvements

Battery-life requirements spurred the innovations and enhancements to the sleep states in the mobile platform. Non-mobile platforms do not have a battery life requirement; instead, the focus there is to reduce overall energy usage. Energy Star requirements in the United States and similar requirements elsewhere in the world specify the maximum energy usage under various idle conditions for compliance. The energy usage in Idle is directly determined by the deepest idle

power sleep state. When the processor is in a deep sleep state, the activity on the platform is very low. Hence, the total power and energy savings on the platform will be equal to the sum of the savings in the processor, the chipset, the VRM, et cetera [5].

The desktop processors that predate the Penryn family of processors used the Autohalt and Stop Grant states for the processor. The Penryn family of processors adds support for Deeper Sleep state in desktop platforms to lower the processor and platform idle power. The processor communicates to the platform (via the Graphics Memory Controller Hub (GMCH) and I/O Controller Hub (ICH)) that it is in the Deeper Sleep state. This allows the GMCH and ICH to power off portions that are required to service the processor requests. This enables further reduction of the platform power when the processor is idle.

The Deeper Sleep state reduces processor power consumption by reducing the voltage of the processor to a lower level. At this level, the processor cannot execute instructions, but its state is retained. When in a sleep state, the power is determined by the amount of leakage in the processor. Leakage is a strong function of the voltage, and hence it is lower in the sleep states in which voltage is reduced. The idle power in the Deeper Sleep state was estimated to be significantly lower than the power in Autohalt or Stop Grant state due to the associated voltage reduction. The target power in the Deeper Sleep state is about 50 percent lower than the power in the Autohalt and Stop Grant states.

The Penryn family of processors' desktop platform also implemented one more platform power-saving

feature along with the Deeper Sleep state. The VRM losses can amount to more than 10 percent of the platform power when in idle state. Most VRMs are optimized to work at medium to highly-loaded conditions. This means that they are less efficient when they are lightly loaded as in the case of the processor being in Deeper Sleep state. The Penryn processor family communicates to the VRM when it enters the Deeper Sleep state, and it lets the VRM shut down all but one phase. This reduces the conversion losses in the other phases and also boosts the efficiency of the single active phase.

Desktop platforms use both the quad-core and dual-core processors of the Penryn family. Support for the Deeper Sleep state is available in both the dual- and quad-core configuration. The challenges of extending Deeper Sleep state support to the quad-core configuration are described in the previous section.

Enabling deeper sleep state on desktops

The Deeper Sleep state was exposed in the ACPI tables [3] as the C3 state. The Deeper Sleep state has a longer latency to memory traffic and interrupt response compared to the Autohalt and Stop Grant state as a result of its having to ramp up the voltage to the processor before responding to memory snoops and interrupts. There was a concern that this increased exit latency could have two undesirable consequences. Firstly, devices that were not designed to tolerate the latency to memory accesses could have buffer overruns and fail. Secondly, specific applications could suffer performance degradation due to the increased interrupt latency. To address the concerns, the exit latency was tuned down to the minimum value possible in the platform timers via a BIOS configuration. Various peripherals such as USB and Firewire (IEEE 1394) were tested for memory access latency increases. Benchmarks such as Sysmark were tested to understand the implications of increased interrupt latency. The results showed that the device functionality was unaffected, and the benchmark score differences were within the normal run-to-run variations. Performance benchmarks such as SPEC* do not have to be tested with Deeper Sleep because the processor is never idle during those benchmark runs.

The conclusion was that introducing Deeper Sleep state to the desktop platform did not have any noticeable adverse impact.

Power constrained performance in multi-core processors

EDAT takes advantage of the power headroom of the idle core to boost the performance of the active core

while running an ST application. The EDAT frequency, which is pre-programmed in the chip, is chosen such that the total power still remains within the specified TDP as illustrated in Figure 3. In the Penryn mobile platforms, this optimization provides a 10 percent frequency boost for ST applications. The Penryn family of processors also implements a hysteresis mechanism that allows it to tolerate short wake-up intervals of the idle core without having to exit the EDAT frequency. This helps minimize performance loss in high-interrupt-rate workloads.

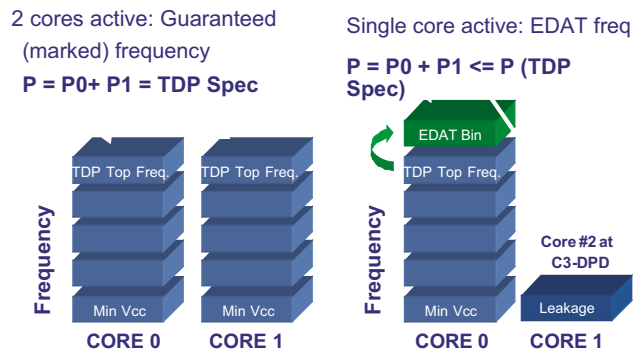


Figure 3: EDAT power budget reallocation.

Multi-core architectures with two, four, or more cores are the standard for power-efficient computing. However, software is lagging in terms of being able to employ all available cores, either because applications have limited threading potential or because tools are not readily available to (re-)write or (re-)build code into threaded applications. As the number of cores within a package grows, many platforms, mobile in particular, become thermally constrained because they are designed to a power target primarily based on form factor. For multi-core designs, this power target usually assumes a workload that pushes the TDP envelope of all the cores. This implies that workloads that do not utilize all the cores underutilize the thermal capacity of the platform. This phenomenon will only grow more prevalent as the computer industry heads towards enabling smaller and cooler form factors.

The Penryn family of processors implemented EDAT in dual-core and dual EDAT in quad-core configurations to intelligently utilize the power headroom from idle cores and opportunistically boost the performance of applications that do not utilize all available cores, without exceeding the system thermal design constraints.

EDAT principles

EDAT operation is based on the following assumptions:

List Item (Number Type: 1,2,3,...)

1. The EDAT frequency (f_{EDAT}) for dual-core and quad-core processors is one frequency bin—typically 267 MHz or 333 MHz—over the maximum TDP-limited (guaranteed) frequency.
2. In dual-core processors, running one core at EDAT frequencies while the other core is idle will not exceed the TDP limit of the processor. Similarly, in quad-core processors, running one or two cores at EDAT frequencies while the other cores are idle will not exceed the TDP limit of the processor.
3. The EDAT frequency is requested by software via the legacy SpeedStep interface that sets the processor's frequency and voltage (F/V) operating point.
4. The processor will transition to f_{EDAT} only if the appropriate number of processors is idle and if the operating system requests the highest performance state (P-state).
5. If there are not enough idle cores available to meet the EDAT criteria when the operating system requests the highest P-state, then the processor will run at the guaranteed operating point.

Assumptions 1 and 2 allow worst-case power assumptions to be applied to running cores during EDAT operation without having to worry about violating TDP limits.

Assumptions 3 and 4 ensure that EDAT will not be activated independently by hardware, without an operating system request, to enter a high-performance state; therefore these assumptions prevent the processor from consuming high power while lightly loaded. These assumptions also ensure that the processor will deliver at least the guaranteed performance level regardless of whether EDAT can be activated.

(Figure 4) depicts how the SpeedStep mechanism normally takes operating system P-state requests (i.e., F/V operating point) and compares them against a fixed guaranteed F/V operating point before determining the processor's "resolved" F/V operating point. If the requested operating point is above the guaranteed operating point, the request is clipped to the guaranteed operating point. It also shows how EDAT dynamically changes the F/V limit between the guaranteed and EDAT limit. The additional logic lets the processor run at the higher EDAT frequency based on how many cores are active, and a hysteresis mechanism ensures the hardware does not switch too often between the guaranteed and EDAT operating points.

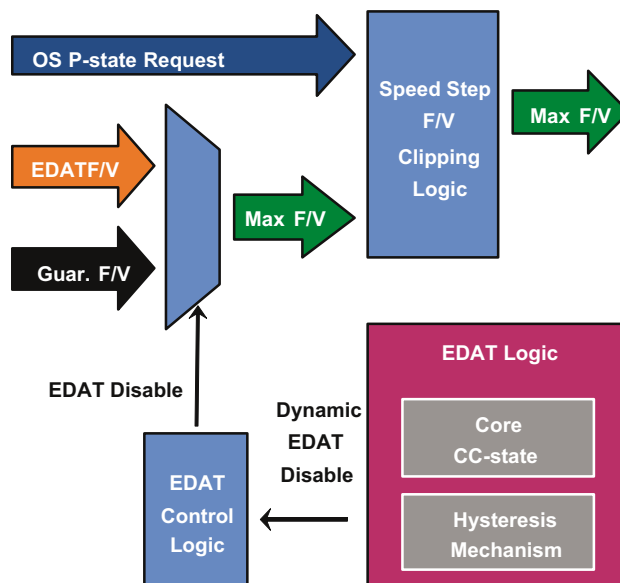


Figure 4: EDAT control mechanism.

Idle cores and wake-up rates

The number of idle cores is one of the inputs to the EDAT decision logic. For the Penryn family of processors, cores are considered idle when their CC state is CC3, CC4, or CC6. In these states, cores consume only leakage power, and their clock distribution networks are shut off, which frees up TDP headroom for any active cores to use to run at f_{EDAT} . Furthermore, dedicated caches/buffers are flushed in these states, which allows any active cores to operate without having to wake up the idle core for cache snooping.

The hysteresis mechanism is the second input into the EDAT decision logic. It permits the processor to run at f_{EDAT} even if the number of idle cores does not meet the minimum requirement for a limited time, and it ensures that the processor stays within power and thermal constraints. This addresses a performance glass jaw that could be encountered if EDAT is simply disabled as soon as idle cores wake up.

Without the hysteresis mechanism, it is possible for idle cores to be frequently woken up by break events but remain active just long enough to service them before going idle again. Service times can be as short as 10 μ s to 20 μ s for Timer Interrupt Service Routines. This can cause rapid transitions in and out of the EDAT operating point and result in performance degradation due to SpeedStep transition overhead and from running the processor at the guaranteed frequency while servicing the break event. A common example of this is when multimedia applications set the timer interrupt rate to 1000 interrupts per second.

Having two cores active for 10 us to 20 us at a time is not likely to cause any thermally significant change on the platform since thermal time constants are in milliseconds. This means that EDAT does not have to be deactivated every time idle cores wake up. The hysteresis mechanism detects these situations and avoids unnecessary SpeedStep transitions.

An example of how the hysteresis mechanism works in a dual-core scenario is illustrated in Figure 5.

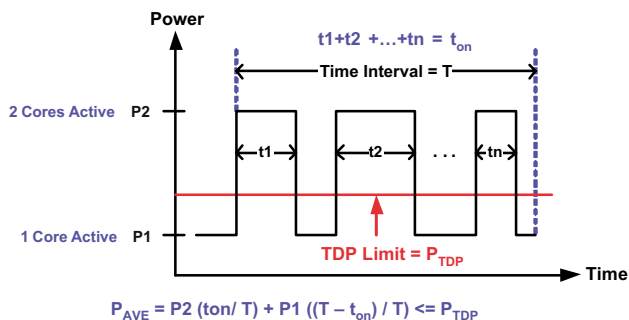


Figure 5: Hysteresis mechanism and average TDP.

- P1 is the worst-case power usage when one core is active at f_{EDAT} , and P2 is the worst-case power usage when two cores are active at f_{EDAT} . P_{TDP} can be arbitrarily chosen to be some small percentage, less than 5 percent, higher than the TDP of a top-bin, non-EDAT processor. In order to allow two

cores to run at f_{EDAT} , platform voltage regulators need to be able to supply the current for two cores active at that frequency for short durations. The cost of supplying the larger current for short durations is usually small and is expected to add an insignificant cost to the platform.

- The time interval, T, is in milliseconds and is based on platform thermal time constants. Using this definition of T, along with the power data above, it is possible to compute the time that two processors can be at f_{EDAT} , t_{on} , such that the average power, P_{AVE} , over the interval, T, is less than or equal to P_{TDP} .
- If two processors are active at f_{EDAT} for more than t_{on} in any interval, T, the hysteresis mechanism “expires” and EDAT is disabled. The hysteresis mechanism also gates re-entering f_{EDAT} to avoid pathological cases in which a core immediately goes into CC3 right after exiting f_{EDAT} , the processor re-enters f_{EDAT} , and the idle core comes out of CC3. In this example, the total time that two cores are active at f_{EDAT} could exceed t_{on} across the two f_{EDAT} periods.

EDAT extension into quad-core architecture

EDAT support in Penryn quad-core processors—dubbed Dual EDAT—was achieved in the same way that we enabled Deeper Sleep State in quad-core processors. The quad-core power-management coordination functionality was extended for Dual EDAT to

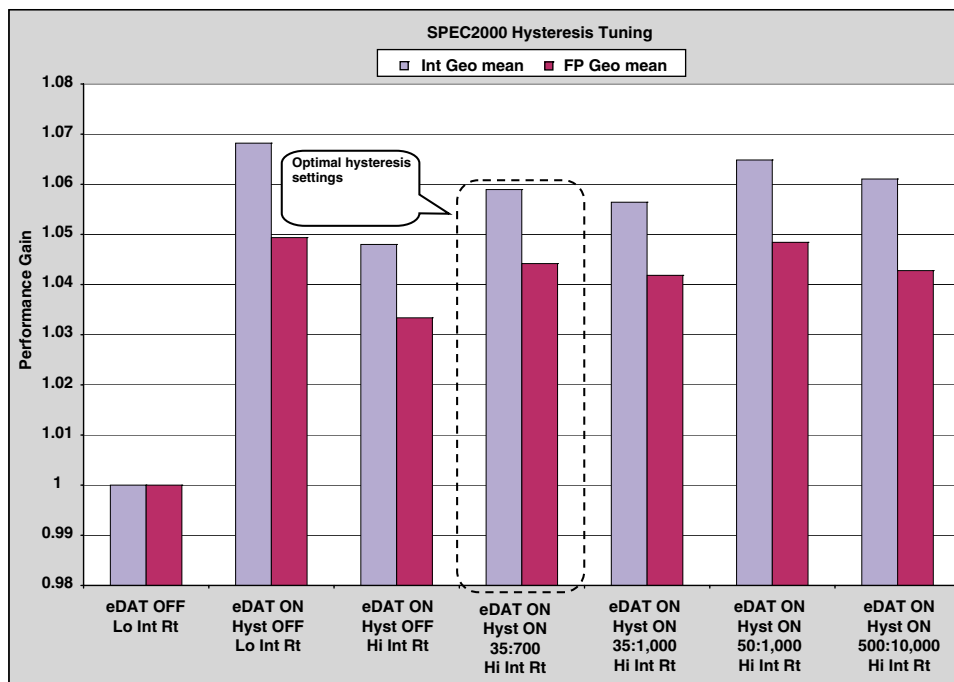


Figure 6: EDAT hysteresis mechanism tuning (Estimated SPEC* CPU2000 as measured on pre-production systems).

share core idleness information and if EDAT was enabled by BIOS and or software. Each site locally resolved whether or not to allow the processor to run at EDAT frequencies, and it was expected that, in steady-state, they would both either allow or disallow the processor to run at f_{EDAT} frequencies.

RESULTS

DPD

Average power on the Penryn mobile processor was benchmarked using the MobileMark® 2005 (MM05) battery life benchmark. The benefit of DPD is shown in Figure 7. These measurements were done on multiple Penryn processor parts with varying leakage power behavior. DPD reduced the average power by approximately 40 percent or more on most of the parts tested. This testing was done on Intel Customer Reference Board platforms using a fresh build of Windows XP* and the MM05 benchmark according to the guidelines specified by BAPCO* [6]. The processor power was measured by sensing the processor voltage and current during the MM05 run and averaging it.

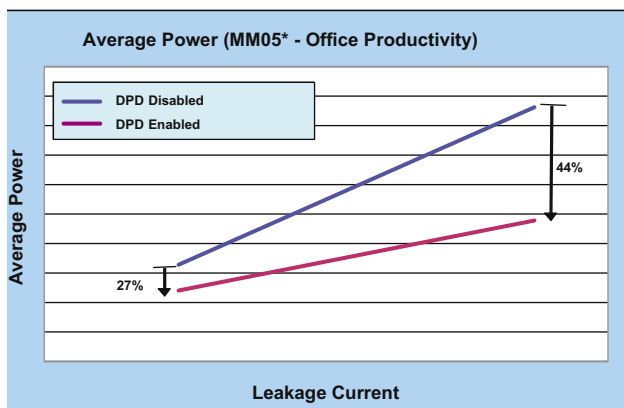


Figure 7: DPD power savings.

Deeper sleep in mobile quad core processors

The same measurements as for DPD (listed above) were done for mobile quad-core processors. The average power on MM05 under Windows XP* was within 15–30 percent of the dual-core processors.

CC3 in servers

CC3 power-saving measurements were done on the Intel Customer Reference Board platform with the Seaburg chipset and ESB2. The dual-processor CPUs based on Intel Xeon® technology, from the Penryn family of processors, at 3.2 GHz with a 1600 MHz Front Side Bus (FSB) had two processors installed in the system for the

tests. The server workloads were applied by SPECpower_ssj2008. The power saving was compared to the case in which only C1E was enabled and CC3 was not enabled. These configuration changes were made using the Intel BIOS options to enable and disable CC3. The savings range from 0 percent at complete idle to a savings of 10–20 percent at medium loads. The savings decrease at maximum loaded conditions as the idle time in C1E and CC3 decreases. This is shown in Figure 8.

Deeper sleep in desktop platforms

The measurements for idle power improvements were measured on the Intel Desktop Customer Reference board with dual-core and quad-core desktop processors from the Penryn family of processors. These measurements were done under operating system idle conditions. The idle power decrease from Autohalt Stop Grant to Deeper Sleep state was in the range of 40 percent–60 percent.

EDAT

Performance measurements on the latest Penryn family of processors mobile quad-core processors showed that Dual EDAT gave the expected performance gains on SPECInt2000, SPECFP2000, and their two-threaded SPEC rate counterparts. (Refer to “45nm Intel Core 2 Processor Silicon Performance” in this issue of the Intel Technology Journal, where Figures 11 and 12 show that Dual EDAT yielded a healthy 5 percent to 8 percent performance increase on these workloads for an 11 percent frequency boost over the guaranteed frequency [4]). Note that Dual EDAT has an idle core limit of two or more; hence, running SPEC rate with three or more copies does not result in any performance improvement.

Like most of the power and performance features, we tuned the EDAT hysteresis mechanism post-silicon to give the best performance on processor-intensive benchmarks such as SPEC2000. Figure 6 indicates that it recovered over half of lost performance gains due to high interrupt rates.

For EDAT hysteresis measurements, we used the Mobile Customer Reference board configuration with a Penryn 2.40 (HFM) 2.60(EDAT) GHz, an 800 MHz Bus, and 2 × 2 GB FB DIMM at 667 MHz. The operating system was Windows XP SP2*. The Workload was SPEC2000 measured by enabling and disabling the EDAT hysteresis mechanism with a timer interrupt rate at 1,000 interrupts second.

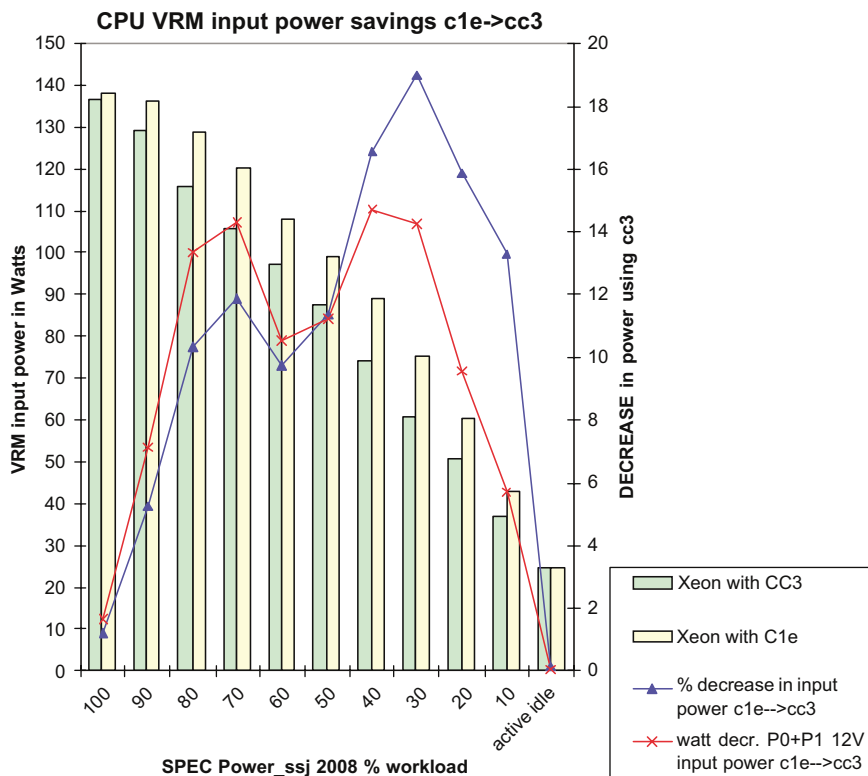


Figure 8: CC3 Power Savings vs. CPU Load

CONCLUSION

All the idle power improvement features introduced in the Penryn family of processors, namely DPD, CC3 in servers, Deeper Sleep in the desktop platform, and quad-core processors, met or exceeded the pre-silicon estimations. To the end user this means an increase in battery life on the mobile platforms, or a decrease in electricity usage in the desktop and server platforms. All these power improvements were achieved without any loss in functionality and negligible performance impact on real-world applications and benchmarks.

EDAT and Dual EDAT provide a means of efficiently utilizing the power envelope of multi-core platforms such that performance is not always constrained by the worst-case TDP scenario. This technology will be employed by this and the next-generation Intel processors.

ACKNOWLEDGEMENTS

Eric Heit and Suresh Doraiswamy were key contributors to the definition and implementation of these power-management features on the Penryn family of processors. We consulted with Steve Fischer and Rob Milstrey who provided guidance throughout the definition, implementation, and productization of these features. Alon Naveh and Efraim Rotem provided help and guidance on the 65nm Intel Core

microarchitecture power and thermal management. Asim Nisar provided performance characterization for the EDAT and Dual-EDAT features. Ivan Herrera, Evgeny Vaisman, and Skip Lindsay were the key validators of the power-management features. Paul Zagacki played a key role in the definition and enabling for Deep Sleep on the desktop platform. Prabhu Rajamani, John Trelford, and Vamsi Jakkampudi helped with silicon data collection and characterization. Susumu Arai, Barnes Cooper, and Anil Aggarwal provided consultation on BIOS, operating systems, and OSPM aspects.

REFERENCES

- [1] <http://www.intel.com/technology/itj/>.
- [2] Intel Technology Journal, Vol. 12, No. 3 2008 at <http://www.intel.com/technology/itj/2008/v12i2/.htm>.
- [3] ACPI Specification at www.acpi.info.
- [4] N. *et al.* 45nm Intel Core 2 Processor Silicon Performance. Intel Technology Journal 2008 Vol. 12, No. 3.
- [5] Paul Z. *et al.* Desktop Platform Power Improvements. Intel Technology Journal, Vol. 12, No. 3, 2008.

[6] BAPCO MobileMark 2005 at <http://www.bapco.com/products/mobilemark2005/>.

AUTHORS' BIOGRAPHIES

Jose Allarey joined Intel in 1998 and is currently working as a Computer Architect in the Mobility Group at the Folsom Design Center. His areas of specialization are Power and Thermal Management. He has six patents pending in this area. He received a B.S. degree in Electrical Engineering from the University of the Philippines in 1992 and an M.S. degree in Computer Architecture from Purdue University in 1998. His email is jose.p.allarey at intel.com.

Varghese George is a Principal Engineer in Intel's Mobility Group. He currently leads the MG-US Architecture team. For the Penryn family of processors, he led the Power Thermal Management Architecture team. Varghese joined Intel in 1993 and has worked on various processor products in areas of microarchitecture, multi-processor performance analysis, and power management. His focus in the last few projects has been on power-management techniques on-chip, and he has been instrumental in architecting many key power-management features into Intel processors. He holds more than 15 patents in various domains and has more pending. Varghese holds a B.S. degree in Electrical Engineering from the University of Mysore in India and an M.S. degree in Computer Engineering from the University of Maryland, College Park. His email is varghese.george at intel.com.

Sanjeev Jahagirdar joined Intel in 1996 and is currently working as a Computer Architect in the Mobility Group at the Folsom Design Center. His areas of specialization are power and thermal management. He holds seven patents in this area and has more pending. He received a B.S. degree in Electrical Engineering from the College of Engineering, Poona, India; in 1992 and an M.S. degree in Computer Architecture from Arizona State University in 1996. His email is sanjeev.jahagirdar at intel.com.

Glossary

C1	Autohalt state
C1E	Enhanced Autohalt state
C3	Deep Sleep state
C3E	Enhanced Deep Sleep state
C4	Deeper Sleep state
C4E	Enhanced Deeper Sleep state
C5	Enhanced Deeper Sleep state
C6	Deep Power Down state
CC	Core level C state. For example, CC3 is a core-level Deep Sleep state

BunnyPeople, Celeron, Celeron Inside, Centrino, Centrino logo, Core Inside, FlashFile, i960, InstantIP, Intel, Intel logo, Intel386, Intel486, Intel740, IntelDX2, IntelDX4, IntelSX2, Intel Core, Intel Inside, Intel Inside logo, Intel. Leap ahead., Intel. Leap ahead. logo, Intel NetBurst, Intel NetMerge, Intel NetStructure, Intel SingleDriver, Intel SpeedStep, Intel StrataFlash, Intel Viiv, Intel vPro, Intel XScale, IPLink, Itanium, Itanium Inside, MCS, MMX, Oplus, OverDrive, PDCharm, Pentium, Pentium Inside, skool, Sound Mark, The Journey Inside, VTune, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Intel's trademarks may be used publicly only with permission from Intel. Fair use of Intel's trademarks in advertising and promotion of Intel products requires proper acknowledgement.

Any codenames featured in this document are used internally within Intel to identify products that are in development and not yet publicly announced for release. For ease of reference, some codenames have been used in this document for products that have already been released. Customers, licensees, and other third parties are not authorized by Intel to use codenames in advertising, promotion or marketing of any product or services and any such use of Intel's internal codenames is at the sole risk of the user.

*Other names and brands may be claimed as the property of others.

**Data from a sample distribution of parts.

Microsoft, Windows, and the Windows logo are trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

Bluetooth is a trademark owned by its proprietor and used by Intel Corporation under license.

SPEC, SPECint and SPECfp are registered trademarks of the Standard Performance Evaluation Corporation. For more information on SPEC benchmarks, please see www.spec.org.

BAPCO and MobileMark are registered trademarks of Business Applications Performance Corporation. For more information please see www.bapco.com.

Intel Corporation uses the Palm OS[®] Ready mark under license from Palm, Inc.

LEED - Leadership in Energy & Environmental Design (LEED[®])

Copyright © 2008 Intel Corporation. All rights reserved.

Additional legal notices at: <http://www.intel.com/sites/corporate/tradmarx.htm>.