Energy Optimal Speed Control of Devices with Discrete Speed Sets* †

Ravishankar Rao and Sarma Vrudhula
NSF Center for Low Power Electronics
Computer Science and Engineering Department
Arizona State University, Tempe, AZ 85281
{ravirao, vrudhula}@asu.edu

ABSTRACT

We obtain analytically, the energy optimal speed profile of a generic multi-speed device with a discrete set of speeds, to execute a given task within a given time. Current implementations of energy efficient speed control policies (including DVFS) almost exclusively use the minimum feasible speed pair, which has been shown before to be suboptimal. Unlike previous works, ours does not require an explicit functional relationship between the device's power and speed (e.g. the CMOS power model), but only assumes that the power-speed relationship is a W-convex (a discrete equivalent of a convex) function. This assumption allowed us to show that the optimal speed profile uses at most two speeds, and that all the essential characteristics of the power-speed relationship can be encapsulated within a single speed, ω_u. The latter speed is intrinsic to the device (i.e. task independent) and can be readily computed from its power-speed values (without any curve fit). Further, ω_u is also the speed at which the the device consumes the least energy per unit work done. The problem formulation reduces to a linear program in the number of supported speeds, which in general, is difficult to solve analytically. However, the optimum solution has a very simple form – it is either ω_u, or the minimum feasible speed pair for the given task. We verified that a number of commercial DVFS processors, and other devices like disk drives satisfied our model of the W-convex power-speed relationship.

Categories and Subject Descriptors

C.4 [**Computer Systems Organization**]: Performance of Systems; B.8 [**Hardware**]: Performance Analysis and Reliability

General Terms

Algorithms, Performance, Experimentation, Theory

Keywords

voltage scaling, frequency scaling, speed control, low-power, convex functions, energy optimization

1. INTRODUCTION

Power management and speed control have been the two popular techniques used for system level energy management. While power mode switching (also known as Dynamic Power Management (DPM)) techniques [3, 11, 19] aim to reduce the idle mode energy consumption, speed control methods (like Dynamic Voltage and Frequency Scaling (DVFS)) try to minimize the active mode energy when peak performance is not required. DPM techniques were initially developed for hard drives, but are now implemented in almost every device used in computing systems like processors, displays, network interfaces, etc.

Similarly, the DVFS technique, originally developed for processors [5, 7], has inspired similar efforts for energy optimization of other devices like the hard drive [17, 6, 22], display [4], network card [23], etc. This trend suggests that it would be desirable to develop a generic model of a device, and find the energy optimal speed control/power management technique for that device. The resulting generic solution could then be applied to a variety of devices and problem domains. This work represents a first step towards developing such a general framework for energy optimization. Most real devices can only support a small discrete set of speeds, and in this paper, we solve the problem of computing the energy optimal speed profile of such a device, for a single task.

The scheduling algorithms proposed by Weiser et al. [24] suggested that the energy optimal policy would scale down the processor speed to the minimum feasible speed (the "just-in-time" principle). For a processor with discrete speed sets, Ishihara and Yasuura [12] showed that if V^* is the optimum voltage assuming a continuous set of speeds is available, then, the two immediate neighbors of V^* in the actual discrete speed set of the processor constitute the optimal speed profile. While the above results are intuitively appealing and form the basis of a large body of work that build on the DVFS technique, they are in general, sub-optimal for modern processors because they don't account for static power components.

Using CMOS models for both dynamic and leakage components, Martin et al. [15] and Jejurikar et al. [13] computed the optimal supply voltage and body bias voltage for a DVFS processor. Their results suggest that the lowest speed is not necessarily energy optimal. More general results can however be obtained by noting that the power-speed relationship is often convex. Irani et al. [10] and Lorch et al. [14] model the power-speed relationship as a convex

*This work was carried out at the National Science Foundation's State/Industry/University Cooperative Research Centers' (NSF-S/IUCRC) Center for Low Power Electronics (CLPE). CLPE is supported by the NSF (Grant EEC-9523338), the State of Arizona, and an industrial consortium. This work was also supported by NSF through grant CCR-0205227. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

†A full version of this paper is available online at `http://veda.eas.asu.edu/papers/rao-dac05.pdf`.

DAC 2005, June 13–17, 2005, Anaheim, California, USA.

function. As convex functions are not defined for discrete speed sets, a different approach is needed that can model not only processors, but other commonly used devices that only have discrete speeds. Based on a CMOS power model, Miyoshi et al. [16] proposed the critical power slope criterion to identify the energy efficient operating points of a DVFS processor with discrete speeds.

Our work is closest to the latter work and that of Ishihaara and Yasuura [12], and makes the following contributions: (i) We derive an *analytical solution* to the single device energy optimal speed control problem. (ii) Our model of a device is sufficiently general that the proposed speed control method can be applied to a *wide range of devices*. (iii) The method is applicable to both discrete and continuous speed sets, accounts for energy overheads, and *does not require a functional form* for the power-speed relationship. (iv) We showed that each device has an *intrinsic speed* ω_u (that depends solely on its power-speed relationship) at which it consumes the least energy per unit work done. (v) The final solution to this general problem has a *very simple form*, but was proved to be optimal after extensive analysis based on the properties of (discrete equivalents of) convex and quasi-convex functions.

2. PROBLEM FORMULATION

A device is a system with an input variable ω called the **speed** and an output variable P called the power. The integral of the power P and the speed ω over a time interval are respectively defined as the **energy** E consumed and the **distance** traversed θ over that period. The distance traversed is a measure of the "work" done by a device towards executing a given task. In a processor, for example, the speed is the clock frequency,[1] and the distance traversed corresponds to the number of clock cycles executed in a given interval while the device is in the active state while for a disk drive it is the number of data units transferred. Similarly, in a disk drive, the angular speed of the spindle motor is the speed while the number of data units transferred in a given interval is the distance.

The device can exist in two states - the active state and the standby state.[2] A transition from the active state to the standby state incurs a **wakeup energy** overhead of E_w, while the reverse transition incurs no overhead. A transition between any two speeds in the active state causes a **speed change energy** overhead of E_c. It is assumed that $E_c \leq E_w$ and that any delays due to the wakeup or speed-change transitions are much smaller than the deadlines of the tasks executed on the device. The device speed in the standby state is zero and useful work can only be done in the active state. In the active state, the speed ω is an element of the **speed set** $\Omega = \{\omega_1, \ldots, \omega_N\}$, where $0 < \omega_1 < \ldots < \omega_N, N \geq 2$ is the number of supported speeds. For convenience, we define now the **index set** K as $\{1, \ldots, N\}$. The left neighbor $l_\Omega(\omega)$ of an arbitrary speed ω is the greatest speed that is less than or equal to ω and lies in Ω. The **right neighbor** $r_\Omega(\omega)$ can be similarly defined. The pair $(l_\Omega(\omega), r_\Omega(\omega))$ are called the **neighboring speeds** of ω.[3]

We observed the relationship between the power and the speed for a number of commercial processors [7, 9, 8, 1] and for a recently developed multi-mode hard drive [17]. Power-speed data for other devices is not available in the public domain, as support for speed control is still at the experimental stage for many of them. For the devices we studied though, the power-speed relation was found to satisfy the following property:

Consider any $\omega_i, \omega_j \in \Omega$ and any $\lambda \in [0,1]$ such that the speed $\omega_k = \omega_i\lambda + \omega_j(1-\lambda)$ belongs to the speed set Ω. Let P_i, P_j and P_k be the device power consumption at speeds ω_i, ω_j and ω_k, respectively. Then, the following inequality holds $P_i\lambda + P_j(1-\lambda) \geq P_k$. Our device model assumes that the power-speed function satisfies this property, and we define such a power-speed function as a **W-convex** function. Figure 1 shows an example of such a function. The above property is a discrete version of a defining property of convex functions – a chord drawn between two points on the curve lies above all points of the curve between those two points. While we could have the assumed that the power-speed relation is simply a "sampled" version of an underlying convex function, such an assumption would be unnecessarily strong and may not be strictly satisfied by existing processors.

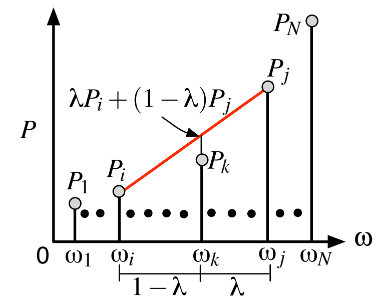

Figure 1: An example of a W-convex function

A task is a specification of the amount of work to be done in a given interval. Formally, a **task** (θ, T) is defined as a requirement that θ units of distance be traversed by the device within a time T. A **speed profile** over an interval $[0,T]$ is a description of the speed $\omega(t)$ at each instant $t \in [0,T]$. Without loss of generality, a speed profile can be represented by the set $\{t_1, \ldots, t_N\}$ where $0 \leq t_k \leq T, k \in K$ is the time for which the device is operated at speed ω_k. The total time spent in the active state is called the **active time** $\tau = \sum_{k\in K} t_k \leq T$, and that in the standby state is $T - \tau$. The **average speed** $\bar{\omega}$ and the **average power** $\bar{P}$ of a speed profile are respectively defined as $\bar{\omega} = \sum_{k\in K} \omega_k t_k / \sum_{k\in K} t_k$, and $\bar{P} = \sum_{k\in K} P_k t_k / \sum_{k\in K} t_k$. The energy consumed by a speed profile X is denoted as E_X. The **minimum average speed** $\bar{\omega}_{\min}(T)$ of the device for a given task (θ, T) is defined as $\bar{\omega}_{\min} = \theta/T$. But, $\tau \leq T$ or $\theta/T \leq \theta/\tau = \bar{\omega}$ so that the average speed of a speed profile must be no less than the minimum average speed for the task.

In general, the energy consumed by an arbitrary speed profile consists of components expended during the active state, standby state, speed change and wakeup. While the latter two components make it difficult to obtain a closed form expression for the total energy, it is useful to first formulate the problem in the absence of such overheads. Then, the energy consumed by an arbitrary speed profile $\{t_1, \ldots, t_N\}$ for a task (θ, T) is $\sum_{k\in K} P_k t_k + P_s(T-\tau) = \sum_{k\in K} (P_k - P_s) t_k + P_s T$. We can see that the last term in this equation is independent of the speed profile. Also, if P_k is W-convex, so is $P_k - P_s$. We assume in the interest of clarity that each of the terms P_k already includes the $-P_s$ in it. The optimization problem for discrete speed sets (in the absence of overheads) can be formally stated as follows: $\min_{t_k, k\in K} P_k t_k$, subject to $\sum_{k\in K} \omega_k t_k / \sum_{k\in K} t_k \geq \bar{\omega}_{\min}$, and $\sum_{k\in K} t_k \leq T$, and $t_k \geq 0 \; \forall \; k \in K$.

The above formulation constitutes a linear program in N decision variables. While it is easy to solve numerically, we believe the insight gained from an analytical solution is worth the extra effort. Further, as the above formulation ignores overheads, the actual energy optimization problem (which we solve in this paper) is more complex than a linear program, because it requires in general, the use of unit step functions to model the overheads.

[1] As the voltage is changed in proportion to frequency, there is essentially a single control variable.

[2] The idle state is a special case of a standby state.

[3] If $r_\Omega(\omega) = \emptyset$, then $\bar{\omega}_{\min} > \omega_N$ and the problem has no solution.

3. ENERGY OPTIMIZATION TECHNIQUE

We now present a set of results that break down the above problem into simpler parts. (Their proofs have been omitted for lack of space.) These results are then combined to obtain the energy optimal speed policy. The following result shows that we only need to consider speed profiles with one or two speeds to find the optimal speed policy.

LEMMA 1. *Consider a speed profile $Y = \{y_1, \dots, y_N\}$ such that at least three of the $y_k, k \in K$ are nonzero, and another profile $X = \{x_1, \dots, x_N\}$ such that at most two $x_k, k \in K$ are nonzero. If both profiles cover the same distance in the same active time, $E_X \le E_Y$.*

Consider then, an arbitrary two-speed profile that employs the speeds ω_a and ω_b for times t_a and t_b, respectively[4] so that the active time $\tau = t_a + t_b$. Let $\lambda = t_a/\tau$, so that $t_b = (1-\lambda)\tau$. Note that as $t_a, t_b \ge 0$, we must have $0 \le \lambda \le 1$. Now, let us define the **energy function** $E(a,b,\lambda), a, b \in K$ to be

$$E(a,b,\lambda) = \frac{P_a\lambda + P_b(1-\lambda)}{\omega_a\lambda + \omega_b(1-\lambda)} = \frac{P_a t_a + P_b t_b}{\theta} \qquad (1)$$

or the energy function is proportional to the objective function of our optimization problem so that it is sufficient to minimize E.

Let us define the device **Q-function** as $Q_k = P_k/\omega_k \ \forall\, k \in K$.

LEMMA 2. *Let $a, b \in K$ and $a < b$ (so that $\omega_a < \omega_b$). If $Q_a \le Q_b$, $E(a,b,\lambda)$ is monotonically non-increasing in λ, otherwise, it is monotonically non-decreasing in λ, for all $0 \le \lambda \le 1$.*

COROLLARY 1. *Let $a, b \in K$ and $a < b$. Then, $E(a,b,\lambda)$ is minimized either at $\lambda = 1$ (if $Q_a \le Q_b$) or at $\lambda = 0$ (if $Q_a \ge Q_b$).*

From (1), it follows that the Q-function value at a particular speed ω_k is equal to the energy function of the single speed profile ω_k. The speed that minimizes the Q-function over the device's speed set is called the **unconstrained optimum speed** ω_u and a policy that uses the single speed ω_u is called the **unconstrained optimum speed policy**.

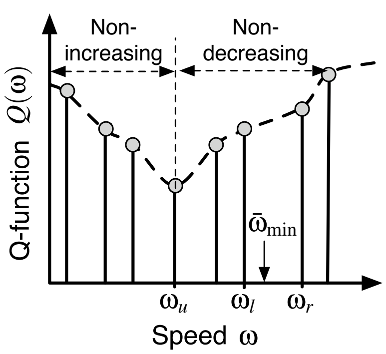

Figure 2: An example of the Q-function

LEMMA 3. *Let $i, j, k \in K$ and $i < j < k$. Then, $Q_j \le \max(Q_i, Q_k)$.*

COROLLARY 2. *Let $i \in K$. Then, Q_i is monotonically non-decreasing in i for $i \ge u$ and non-increasing in i for $i \le u$.*

The above result shows that the Q-function is monotonic on either side of the unconstrained optimum speed[5] (see Figure 2).

LEMMA 4. *If the average speed of a device's speed profile was constrained to equal a given value $\bar{\omega}$ that is greater than the minimum feasible speed $\bar{\omega}_{min}$, the optimum speed profile would be a convex combination of the two neighboring speeds of $\bar{\omega}$.*

[4]A one speed policy is a particular case with either t_a or t_b set to 0.

[5]A property shared by quasi-convex functions [2].

The following lemma deals with the tradeoffs between active energy, standby energy and the energy overheads.

LEMMA 5. *Consider a two-speed profile X whose average speed equals $\bar{\omega}_{min} = \theta/T$ (no wakeup overhead but with a speed-change overhead), and a one-speed profile Y with speed $\omega_k > \bar{\omega}_{min}$ (with wakeup but no speed-change overhead). Let $\bar{P}_X$ be the average power of X, and Q_k be the device Q-function value for speed ω_k. Then $E_X \le E_Y$ if and only if $\bar{\omega}_{min} > \omega_u$ or $\bar{P}_X T + E_c \le \theta Q_k + E_w$.*

The next result now shows that the optimal value of $\bar{\omega}$ is simply the minimum average speed $\bar{\omega}_{\min}$.

LEMMA 6. *If $\bar{\omega}_{min} > \omega_u$, the optimum speed profile chooses its speeds such that the average speed $\bar{\omega} = \bar{\omega}_{min}$.*

Thus, the optimal speed profile chooses the two speeds $l_\Omega(\bar{\omega}_{\min})$ and $r_\Omega(\bar{\omega}_{\min})$ that are the neighbors of the minimum average speed, so that λ^*, the proportion of the active time that speed $l_\Omega(\bar{\omega}_{\min})$ is used for is given by

$$\lambda^* = \frac{r_\Omega(\bar{\omega}_{\min}) - \bar{\omega}_{\min}}{r_\Omega(\bar{\omega}_{\min}) - l_\Omega(\bar{\omega}_{\min})} \qquad (2)$$

Such a speed policy is called the **minimum feasible speed policy**. The active energy consumption of the latter policy is given by $P_l\lambda^* T + P_r(1-\lambda^*)T$, while that of the unconstrained optimum speed policy is given by $P_u\tau = P_u\theta/\omega_u = Q_u\theta$. Defining the difference in the active energies as

$$A(\theta,T) = P_l\frac{\omega_r - \theta/T}{\omega_r - \omega_l}T + P_r\frac{\theta/T - \omega_l}{\omega_r - \omega_l}T - Q_u\theta, \qquad (3)$$

the energy optimal speed policy is given by the following theorem.

THEOREM 1. *If $\bar{\omega}_{min}(\theta,T) > \omega_u$ or $A(\theta,T) \le E_w - E_c$, the energy optimal speed policy is the minimum feasible speed policy, otherwise, it is the unconstrained optimum speed policy.*

This result shows that if the unconstrained optimum speed policy is infeasible, the optimum policy is simply the minimum feasible speed policy. Even if the unconstrained optimum speed policy is feasible though, it is optimal only if the improvement in the active energy A over the minimum speed policy is sufficiently large to offset the additional overhead $E_w - E_c$ it incurs. In the absence of overheads the optimum speed policy simply reduces to the following: if $\bar{\omega}_{\min} > \omega_u$, use the minimum feasible speed policy, else, use the unconstrained optimum speed policy.

4. RESULTS FOR PROCESSOR DVFS

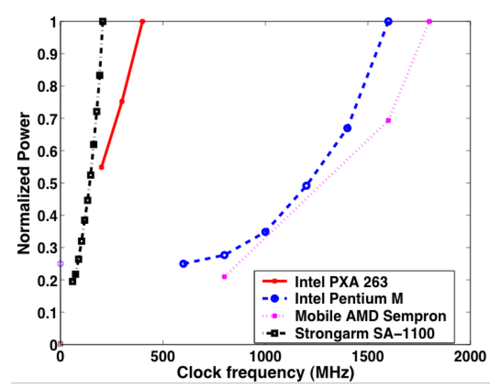

Figure 3: Normalized processor power plots

Due to the lack of space, we have omitted the results for disk drives. We obtained the power-frequency data for four commercial processors supporting either DVFS or Dynamic Frequency Scaling (DFS): Intel PXA 263 [9], Intel Pentium M [8], Mobile AMD Sempron [1] and the Intel Strongarm SA100 [18] (Figure 3). All four

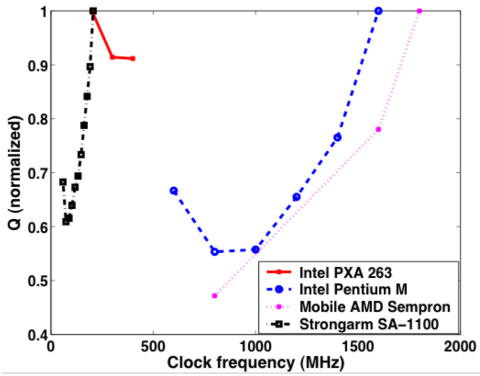

Figure 4: Normalized processor Q-function plots

power functions were found to be W-convex. Figure 4 shows the Q-functions for these processors. From the figure, we can see that the unconstrained optimum speeds for the PXA, Pentium-M, Sempron and Strongarm are 400 MHz (highest frequency), 800 MHz (second smallest frequency), 800 MHz (lowest frequency), 73.7 MHz (second lowest frequency), respectively. For the three processors except the Sempron, we can expect energy improvements over the current speed policy because their unconstrained optimum exceeds their least supported speed. We compare the energy consumption

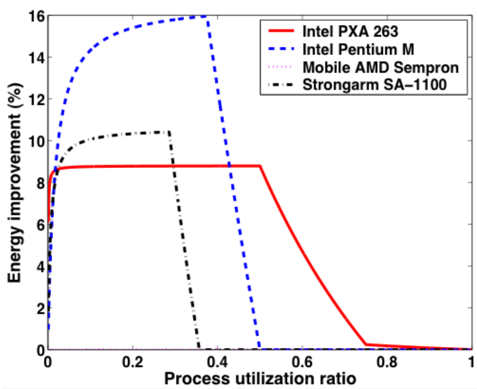

Figure 5: Energy improvement over current DVFS policy

due to the proposed policy with the current policy for DVFS, which is to use the current policy simply uses the minimum speed policy for all values of $\bar{\omega}_{\min}$. For a generic workload expressed in terms of processor utilization, energy improvements of up to 9%, 16% and 10% were observed for the PXA, Pentium M and the Strongarm, respectively (Figure 5). As expected, the Sempron shows no improvements. Not all processor manufacturers provide information about the power-frequency values in their data sheets - higher energy improvements than those shown above may indeed be possible for several commercial processors. These results show we can do better than the current practice of just-in-time speed scaling, and this only requires power-speed data for the processor.

5. CONCLUSION

An analytical solution was obtained for the problem of finding the energy optimal speed profile of a generic device executing a single given task. The solution chooses either of two simple speed policies based on the energy overheads of the device and the relative magnitudes of its unconstrained optimum speed ω_u (found by measurements/data sheets) and the minimum feasible speed $\omega_{\min}$ of the given task (which can be expressed analytically in terms of the task parameters). Unlike previous solutions to similar problems (DVFS in particular), the proposed method is applicable to any device with a W-convex power-speed relationship $P(\omega)$ and does not require a closed form expression for $P(\omega)$. The proposed approach has the following consequences for dynamic energy optimization or energy-aware system design: (i)A DVFS chip can be designed to find its optimum speed by performing a few power measurements for known benchmarks. This general method would work then, for any chip, *regardless of manufacturing technology or process variations.* (ii) The simple analytical and nature of the solution can be used to solve more complex problems. For example, we have been able to use the proposed framework to find the energy optimal speed profile for a pair of generic interacting devices [21]. We intend to extend the proposed speed control methodology further for networks of interacting components, and for battery lifetime optimization (instead of just energy minimization).

6. REFERENCES

[1] Advanced Micro Devices. *AMD Athlon 64 Processor Power and Thermal Data Sheet*.

[2] M. S. Bazaara, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, second edition, 1993.

[3] L. Benini, A. Bogliolo, G. A. Paleologo, and G. De Micheli. Policy optimization for dynamic power management. *IEEE Trans. CAD*, 18(6):813–833, June 1999.

[4] N. Chang, I. Choi, and H. Shim. DLS: Dynamic backlight luminance scaling of liquid crystal display. *IEEE Trans. VLSI Sys.*, 12(8):837–846, August 2004.

[5] M. Fleischmann. Longrun power management^{TM}: Dynamic power management for Crusoe^{TM}processors.

[6] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke. Reducing disk power consumption in servers with DRPM. *IEEE Computer*, 36(12):59–66, December 2003.

[7] Intel Corp. *Enhanced Intel SpeedStep Technology for the Intel Pentium M Processor*.

[8] Intel Corp. *Intel Pentium M Processor on 90nm Process with 2-MB L2 Cache*.

[9] Intel Corp. *Intel PXA26x Processor Family : Electrical, Mechanical, and Thermal Specification*.

[10] S. Irani, S. Shukla, and R. Gupta. Algorithms for power savings. In *Proc. ACM-SIAM Symposium on Discrete Algorithms*, pages 37–46, Philadelphia, PA, USA, 2003. Society for Industrial and Applied Mathematics.

[11] S. Irani, S. Shukla, and R. Gupta. Online strategies for dynamic power management in systems with multiple power-saving states. *ACM Trans. Embedded Computing Sys. (TECS)*, 2(3):325–346, 2003.

[12] T. Ishihara and H. Yasuura. Voltage scheduling problem for dynamically variable voltage processors. In *Proc. Intl' Symp. Low Power Electronics and Design (ISLPED)*, pages 197–202, 1998.

[13] R. Jejurikar, C. Pereira, and R. Gupta. Leakage aware dynamic voltage scaling for real-time embedded systems. In *Proc. Design Automation Conf. (DAC)*, pages 275–280, 2004.

[14] J. R. Lorch and A. J. Smith. PACE: A new approach to dynamic voltage scaling. *IEEE Trans. Computers*, 53(7):856–869, July 2004.

[15] S. M. Martin, K. Flautner, T. Mudge, and D. Blaauw. Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads. In *Proc. ICCAD*, pages 721–725, 2002.

[16] A. Miyoshi, C. Lefurgy, E. V. Hensbergen, R. Rajamony, and R. Rajkumar. Critical power slope: Understanding the runtime effects of frequency scaling. In *Proc. Intl' Conf. Supercomputing (ICS)*, pages 35–44, 2002.

[17] K. Okada, N. Kojima, and K. Yamashita. A novel drive architecture of HDD: "multimode hard disc drive". In *Proc. Intl' Conf. Consumer Electroncis (ICCE)*, pages 92–93. IEEE Press, 2000.

[18] J. Pouwelse, K. Langendoen, and H. Sips. Application-directed voltage scaling. *IEEE Trans. VLSI Sys.*, 11(5):812–826, October 2003.

[19] Q. Qiu, Q. Wu, and M. Pedram. Stochastic modeling of a power-managed system-construction and optimization. *IEEE Trans. CAD*, 20(10):1200–1217, October 2001.

[20] R. Rao. Energy optimal speed control for components of portable systems. Master's thesis, University of Arizona, Tucson, 2004.

[21] R. Rao and S. Vrudhula. Energy optimization for a two-device data flow chain. In *Proc. Intl' Conf. Computer-Aided Design (ICCAD)*, pages 268–274, November 2004.

[22] R. Rao, S. Vrudhula, and M. S. Krishnan. Disk drive energy optimization for audio-video applications. In *Proc. Conf. Compilers, Arch., Synth. Emb. Sys. (CASES)*, pages 93–103, September 2004.

[23] C. Schurgers, O. Aberthorne, and M. Srivastava. Modulation scaling for energy aware communication systems. In *Proc. Intl' Symp. Low Power Electronics and Design (ISLPED)*, pages 96–99. ACM Press, 2001.

[24] M. Weiser, B. Welch, A. Demers, and S. Shenker. Scheduling for reduced CPU energy. In *Proc. Symp. Operating Sys. Design and Implementation (OSDI)*, pages 13–23, 1994.