

# Online Strategies for Dynamic Power Management in Systems with Multiple Power-Saving States

SANDY IRANI, SANDEEP SHUKLA, and RAJESH GUPTA  
University of California at Irvine

---

Online dynamic power management (DPM) strategies refer to strategies that attempt to make power-mode-related decisions based on information available at runtime. In making such decisions, these strategies do not depend upon information of future behavior of the system, or any a priori knowledge of the input characteristics. In this paper, we present online strategies, and evaluate them based on a measure called the competitive ratio that enables a quantitative analysis of the performance of online strategies. All earlier approaches (online or predictive) have been limited to systems with two power-saving states (e.g., idle and shutdown). The only earlier approaches that handled multiple power-saving states were based on stochastic optimization. This paper provides a theoretical basis for the analysis of DPM strategies for systems with multiple power-down states, without resorting to such complex approaches. We show how a relatively simple “online learning” scheme can be used to improve the competitive ratio over deterministic strategies using the notion of “probability-based” online DPM strategies. Experimental results show that the algorithm presented here attains the best competitive ratio in comparison with other known predictive DPM algorithms. The other algorithms that come close to matching its performance in power suffer at least an additional 40% wake-up latency on average. Meanwhile, the algorithms that have comparable latency to our methods use at least 25% more power on average.

Categories and Subject Descriptors: C.4 [**Computer Systems Organization**]: Performance of Systems

General Terms: Algorithms, Performance

Additional Key Words and Phrases: Dynamic, power management, online algorithms

---

## 1. INTRODUCTION

Power management in embedded computing systems is achieved by actively changing the power consumption profile of the system by putting its components into power/energy states sufficient to meeting functionality requirements. For example, an idling component (such as a disk drive) can be put into a slowdown or shutdown state. Of course, bringing such a component back to the active

---

This work was supported by NSF grant CCR-0098335, SRC, and DARPA/ITO supported PADS project under the PAC/C program. In addition, the first author is partially supported by NSF grant CCR-0105498 and by ONR Award N00014-00-1-0617.

Authors' address: Information and Computer Science Department, University of California at Irvine, Irvine, CA 92697; email: {irani;skshukla;rgupta}@ics.uci.edu.

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 2003 ACM 1539-9087/03/0800-0325 \$5.00

state may require additional energy and/or latency to service an incoming task. The input to the problem we consider here is the length of an upcoming idle period, and the decision to be made is whether to transition to a lower power dissipation state while the system is idle. There are several issues in coming to this decision intelligently. For instance, immediate shutdown—that is, shutdown as soon as an idle period is detected—may not save overall energy if the idle period is so short that the powering-up costs are greater than the energy saved in the sleep state. On the other hand, waiting too long to power down may not achieve the best energy reductions possible. Thus, there exists a need for effective (and efficient) decision procedures to manage power consumption. Dynamic power management (DPM) attempts to make such decisions (usually under the control of the operating system) at runtime based on the dynamically changing system state, functionality, and timing requirements [Benini and De Micheli 1998; Benini et al. 2001; Chung et al. 1999a; Hwang et al. 1996; Irani et al. 2002; Ramanathan et al. 2000; Shukla and Gupta 2001; Srivastava et al. 1996].

In a survey of DPM techniques in Benini et al. [2000], the authors classify DPM strategies into two main groups: (a) *predictive schemes* and (b) *stochastic optimum control schemes*. Predictive schemes attempt to predict the timing of future input to the system and schedule shutdown (usually to a single lower power state) based on these predictions. Stochastic optimum control is a well-researched area [Benini et al. 1999; 2000; Chung et al. 1999a; Qiu and Pedram 1999; Qiu et al. 1999; Simunic et al. 1999]. The chief characteristic of these approaches is the construction (and validation) of a mathematical model of the system that lends itself to a formulation of a stochastic optimization problem. Then strategies to guide the system power profile are devised that achieve the most power savings in presence of the uncertainty related to system inputs. While several useful and practical techniques have been developed using the predictive and stochastic optimum control schemes, it is difficult to develop bounds on the quality of the results without extensive simulations and/or model justification.

## 2. COMPETITIVE ANALYSIS AND ITS LIMITATIONS

We approach DPM as an inherently “online” problem in that an algorithm governing power management must make decisions about the expenditure of resources before all the input to the system is available [Borodin and El-Yaniv 1998]. Specifically, the algorithm does not learn the length of an idle period until the moment that it ends. Analytical solutions to such online problems are often best characterized in terms of a *competitive ratio* [Phillips and Westbrook 1999] that compares the cost of an online algorithm to the optimal offline solution that knows the input in advance. Note that the offline algorithm is not a realistic algorithm, since it knows the length of the idle period in advance. It is just used as a point of comparison to quantify the performance loss due to the fact that a DPM strategy must make decisions with only partial information. To be specific, we say that an algorithm is  $c$ -competitive if, for any input, the cost of the online algorithm is bounded by  $c$  times the cost of the optimal offline

Table I. Values for the Power Dissipation and Start-up Energy for the IBM Mobile Hard-Drive

State	Power Consumption (W)	Start-up Energy (J)	Transition Time to Active
Sleep	0	4.75	5 s
Stand-by	0.2	1.575	1.5 s
Idle	0.9	0.56	40 ms
Active	1.9	0	0

algorithm for that input. The offline algorithm has access to the entire input before committing to any decisions. For DPM, the cost of an algorithm is the total amount of energy consumed. The *competitive ratio* (CR) of an algorithm is the infimum over all  $c$  such that the algorithm is  $c$ -competitive.

Competitive analysis [Borodin and El-Yaniv 1998]—that is, analysis to determine bounds on the competitive ratio—can be done either as a case analysis of the various adversarial scenarios [Karlin et al. 1990; Ramanathan 2000] or through theorem-proving or automatic model checking [Shukla and Gupta 2001]. Competitive analysis has proven to be a powerful tool in providing a guarantee on the performance of an algorithm for any input. Competitive analysis is valuable in situations where it is impractical to obtain and process information for predicting future inputs. It also provides an assurance for the designers of the algorithm about the worst possible behavior of an online algorithm. For example, if an online strategy has a competitive ratio of 2, which assures the designer that no matter what input sequence is provided to the strategy, it will never cost the strategy more than twice the cost incurred by the best possible strategy.

There are two chief limitations of the competitive analysis approach to DPM that we seek to address in this paper. The first is that the results tend to be overly pessimistic due to the fact that they examine worst-case behavior. Thus, there is a need to refine the CR bounds that take into account typical input behavior. In many applications, there is a structure in the input sequence that can be utilized to fine-tune online strategies and improve their performance.

The second problem with our earlier work on CR-based DPM strategies [Karlin et al. 1990; Ramanathan et al. 2000, 2002] is that the system model consists of devices with only two power-saving states, namely an idle state and a shutdown state. This is a limitation of most predictive strategies as well [Benini et al. 2000]. In contrast, most real systems consist of components with multiple power states. There are two kinds of systems here: systems with multiple shutdown states as well as systems with multiple operating states. For instance, Table I shows the power states of a portable hard drive [IBM 1996] that consists of three low-power states in which a device can be when it is not processing any requests. Radio modems represent devices that can often be in multiple active states (for instance, transmitting data at various power levels to the radio transmitter). We do not address the problem of choosing among different operating states in this paper.

For systems with more than one device, our protocols can be applied to each device separately. Each device experiences periods of idle time that may or may not coincide with the idle periods experienced by other devices in the system.

A power management strategy can then be applied to a particular device during the idle periods experienced by that device. There will be some dependence between performances for different devices in cases where a request requires more than one device to be satisfied. In these situations, a device that is in an active state may experience some latency because it has to wait for another device to transition from a low-power state to the active state before either can begin work. In this paper, we focus on performance for a single device, although this phenomenon would be an interesting direction for future work.

## 2.1 Related Work

Many strategies for dynamic management of power and energy have been proposed and studied. These include predictive strategies [Hwang et al. 1996; Srivastava et al. 1996], stochastic-modeling-based strategies [Benini et al. 1999; Qiu and Pedram 1999], session clustering and prediction strategies [Lu and De Micheli 1999], online strategies [Ramanathan et al. 2000], and adaptive-learning-based strategies [Chung et al. 1999b]. A survey of DPM strategies that covers the classes of algorithms in predictive and stochastic optimum control categories can be found in Benini et al. [2000]. Lu et al. [2000] present a quantitative comparison of various existing management strategies.

Many predictive dynamic power management strategies [Benini et al. 1999; Chung et al. 1999a; Hwang et al. 1996; Karlin et al. 1990; Lu and De Micheli 1999; Ramanathan et al. 2000; Srivastava et al. 1996] use a sequence of past idle period lengths to predict the length of the next idle period. These strategies typically describe their prediction for the next idle period with a single value. Given this prediction, they transition to the power state that is optimal for the predicted idle period length. In case the prediction is wrong, they transition to the lowest power state if the idle period extends beyond a fixed threshold value. For the sake of comparison with other approaches, we call these predictive DPM schemes *single-value prediction* (SVP) schemes. Of particular interest is the work of Chung et al. [1999b] that addresses multiple idle state systems using a prediction scheme based on adaptive learning trees. Their method has shown an impressively high hit ratio in its prediction.

In the past, stochastic-control-based methods as in Benini et al [1999; 2000]; Simunic et al. [1999]; Chung et al. [1999a]; Qiu and Pedram [1999]; and Qiu et al. [1999] have been limited because they make assumptions about the characteristic probability distribution of the input/job arrivals, the service time distribution of the device, and so on. Although the policies obtained are optimal given these assumptions, there are no guarantees that the assumptions will always hold. These problems have been addressed to some extent in Chung et al. [1999b]; and Lu and De Micheli [1999]. Recently the stochastic-modeling approach has been extended so that the assumption that interarrival times are exponentially distributed can be removed [Glynn et al. 2000]. Stochastic modeling has also been extended to systems which can be modeled by petri nets [Qiu et al. 2000]. One drawback of these approaches is that they require solving computationally expensive optimization problems to derive the optimal solutions. If the usage patterns for the device change, the algorithm has to be

reoptimized to adjust to the new behavior. In Chung et al. [1999a], an adaptive technique has been proposed which tries to overcome this limitation. The authors use sliding windows to keep a long enough history of the input arrivals, and estimate the parameters of the arrival Markov processes. However, this method is complex because they have to solve the optimization problem to find the optimal strategy in a clock-driven or event-driven manner.

There are also a few existing results on competitive analysis of DPM strategies in the online algorithms literature. Some of this work does not address DPM explicitly, but the general framework applies to DPM as well. An example of this is the now classical result described in Phillips and Westbrook [1999] that shows that 2 is the best competitive ratio achievable by any deterministic online algorithm. Probabilistic analysis for online DPM algorithms in two-state systems has been given in Karlin et al. [1990] and Keshav et al. [1995]. They assume that the distribution over the upcoming idle period is known and optimize the algorithm based on that distribution. They give a method to determine the best online algorithm, given a distribution, and show that for any distribution the expected cost of this online algorithm is within a factor of  $e/(e-1) \approx 1.58$  of the expected cost of the optimal offline algorithm. This result is tight in that there is a distribution for which the ratio is exactly  $e/(e-1)$ , although for some distributions the ratio may be less.

## 2.2 Our Contributions and Paper Organization

In Section 4 we present a deterministic algorithm for DPM for multistate devices and show that it is 2-competitive. This result, which extends the earlier analysis described in Phillips and Westbrook [1999], is tight in the sense that there is no constant  $c < 2$  such that there is a deterministic  $c$ -competitive algorithm that works for all multiple power-down state devices. However, it may be possible to have a competitive ratio less than 2 for a specific device, depending on the parameters of the device (e.g., number of states, power dissipation rates, start-up costs, and so on). Note that this deterministic algorithm focuses on a single idle period, and does not depend on any history of previous idle periods. Thus, it is nonadaptive in that it does not use any information about previous idle periods to tune its behavior.

As we have argued above, competitive analysis often yields unduly pessimistic results. Indeed, our goal is to devise a DPM strategy whose total energy expenditure is much lower than twice the optimal offline algorithm. We show that by keeping track of a sequence of past idle periods we can improve the performance of the 2-competitive deterministic strategy in practice by adapting the behavior of the strategy to patterns observed in the input sequence. We do this in Section 5 by modeling the future input by a probability distribution that is learned from the recent history of the length of idle periods. In doing so, we avoid some of the problems with many of the previous strategies mentioned above.

One of the chief limitations of a single-valued prediction (SVP) approach is that it fails to capture uncertainty in the prediction for the upcoming idle period length. For example, if a very short idle period and a very long idle

period are equally likely, these methods are forced to pick a single prediction and pay a penalty in the likely event that the prediction is wrong. Using a probability distribution to model the length of the upcoming idle period allows for a much richer prediction, so that the algorithm can optimize in a way that takes the nature of this additional information into account. Furthermore, we make no assumptions about the form of the distribution governing idle period length, which means that our method can automatically adapt to a variety of applications, and also applies to nonstationary input arrivals.

In Section 5.1 we derive analytical bounds for an online algorithm for DPM that knows the probability distribution governing the length of the upcoming idle period. Section 5.2 provides a systematic means of dynamically constructing an estimate for the probability distribution based on online observations and combining it with the algorithm from the previous section to build an online power management strategy. Experimental results in Section 7 demonstrate the utility of our strategy in balancing power usage with latency.

### 3. SYSTEM MODEL

Consider a device that can be in one of the  $k + 1$  power states denoted by  $\{s_0, \dots, s_k\}$ . The power consumption for state  $i$  is denoted by  $\alpha_i$ . The states are ordered so that  $\alpha_i > \alpha_j$  as long as  $i < j$ . Thus, state  $s_1$  is the *active* state that is the highest power consumption state. Any strategy for an individual idle period can be described by a sequence of thresholds, each of which is associated with a power consumption state. As soon as the idle periods extend beyond a given threshold, the device transitions to the associated power consumption state.

From the manufacturer's specification, we are also given the transition power  $p_{ij}$ , and transition times  $t_{ij}$ , to move from state  $s_i$  to  $s_j$ . Usually, the energy needed and time spent to go from a higher (power) state to a lower (power) state is negligible, whereas the converse is not true. Thus, we simplify the model by considering only the time and power necessary to power up the system. Furthermore, all of the algorithms considered in this paper have the property that they only transition to the active state when powering up and never transition to an intermediate higher-powered state. As a result, we only need the time and total energy consumed in transitioning up from each state  $i$  to the active state (state 0). The total energy used in transitioning from state  $i$  to the active state is denoted by  $\beta_i$ .

We note that in cases where the time and energy cost incurred in transitioning to lower-power consumption states is nonnegligible, they can be easily incorporated by folding them into the corresponding power-up parameters. This can be done as long as the time and energy used in transitioning down is additive. That is, we require that for  $i < j < k$ , the cost to go from  $i$  to  $j$  and then from  $j$  to  $k$  is the same as the cost of going from  $i$  directly down to  $k$ .

The input to the DPM algorithm is simply a sequence of idle periods. For each idle period, the DPM is notified when the idle period begins and then again when the idle period ends. This is the only information required by our DPM. In our experiments the idle periods are derived by our simulator, which receives a time-stamped sequence of requests for service. With each request,

the simulator is told the time of its arrival and the length of time it will take to satisfy the request. If the device is busy when a new request arrives, it enters a queue and is served on a first-come-first-serve basis. In this case, there is no idle period and the device remains active through the time that the request is finished. This means that the number of idle periods is generally less than the number of requests serviced. Whenever a request terminates and there are no outstanding requests waiting in the system, an idle period begins. In these situations, the DPM is invoked to determine which power consumption states the device should transition to and at what times.

If the device is not busy when a new request arrives, it will immediately transition to the active state to serve the new request if it is not already there. In the case where the device is not already in the active state, the request can not be serviced immediately, but will have to incur some latency in waiting for the transition to complete. This delay will cause future idle periods to be shorter. In fact, if a request is delayed, some idle periods may disappear. Thus, we have an interesting situation where the behavior of the algorithm affects future inputs (idle period lengths) given to the algorithm.

Note also that delaying servicing a request will tend to result in lower power usage. For instance, consider the extreme case where the power manager remains in the deepest sleep state while it waits for all the requests to arrive and then processes them all consecutively. This extreme case is not allowed in our model, since we require that the strategy transition to the active state and begin to work on a request as soon as one appears. However, it illustrates the natural trade-off that occurs between power consumption and latency. Our experimental results explore this trade-off for the set of algorithms studied. A more extensive discussion of this trade-off is presented in Ramanathan et al. [2000].

In the next two sections we present our analysis for the deterministic and probability-based algorithms. In this analysis, we do not take into account the delay incurred in returning to the active state because the analysis focuses on an individual idle period. These sections address the problem of minimizing total energy expenditure, given that the length of the upcoming period is arbitrary (deterministic case) or is governed by a fixed probability distribution (probability-based case). The empirical evaluations incorporate the effect of start-up latency as an entire sequence of requests for service arrive through time.

#### 4. THE DETERMINISTIC ONLINE ALGORITHM

We now present our deterministic algorithm for online DPM. The basic idea is that the online algorithm tries to mimic the behavior of the optimal offline algorithm. At each point in time  $t$ , the algorithm is in the state that the optimal algorithm would have chosen if the length of the idle period were exactly  $t$ .

To get the optimal cost, we plot each line  $c = \alpha_i t + \beta_i$ . This is the cost of spending the entire interval in state  $i$  as a function of  $t$ , the length of the interval. Take the lower envelope of all of these lines. Let us call this function,  $LE(t)$ . The optimal cost for an interval of length  $t$  is  $LE(t) = \min_i \{\alpha_i t + \beta_i\}$ . The

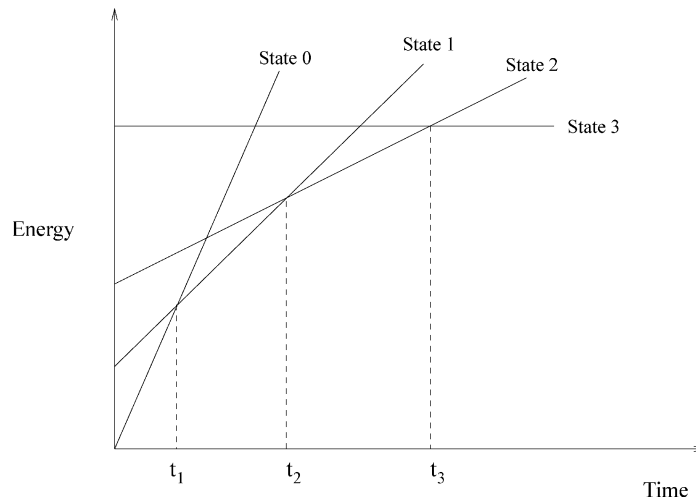


Fig. 1. Energy consumption for each state for a four-state system. Each state is represented by a line that indicates the energy used if an algorithm stays in that state as a function of the length of the idle period. For each state, the slope is the power dissipation rate and the  $y$ -intercept is the energy required to power-up from that state.

online algorithm—called the lower envelope algorithm (LEA)—will also follow the function  $LE$ . It will remain in the state that realizes the minimum in  $LE$  and will transition at the discontinuities of the curve. That is, LEA will remain in state  $j$  as long as  $\alpha_j t + \beta_j = \min_i \{\alpha_i t + \beta_i\}$ , for the current time  $t$ .

For  $j = 1$  to  $k$ , let  $t_j$  be the solution to the equation  $\alpha_j t + \beta_j = \alpha_{j-1} t + \beta_{j-1}$ , where  $t_j$  is the time that LEA will transition from state  $j - 1$  to state  $j$ . We will assume here that we have thrown out all the lines that do not appear on the lower envelope at some point. This is equivalent to the assumption that  $t_1 < t_2 < \dots < t_{k-1} < t_k$ . (See Figure 1).

**THEOREM 1.** *The lower envelope algorithm is 2-competitive.*

The proof of Theorem 1 is shown in the Appendix.

As emphasized earlier, this algorithm does not take into account input patterns. The worst-case scenario, obtained via Theorem 1, shows that the energy cost resulting from the online decisions can be no worse than two times the energy cost of the optimal offline strategy, which knows the input sequence in advance. We show later that depending on request arrival patterns, this worst-case bound may not really happen, and the empirical ratio of the online to offline costs may in fact be much lower. As shown in Qiu and Pedram [1999], Chung et al. [1999b], and Benini et al. [1999], input sequences are often interrelated, and hence modeling of the input pattern and exploiting that knowledge in the design of the algorithm can help bridge the gap between the performance of online strategy and that of the optimal offline strategy. In the next section, we discuss our probability-based algorithm and show that if the probability distribution governing the length of the idle period is known before hand, the worst-case competitive ratio can be improved by 21%, with respect to the



deterministic case. Moreover, we show through experimental evaluation that this worst-case bound is pathological. In fact, we can bring the energy cost of the online algorithm within 27% of the optimal offline one.

## 5. THE PROBABILITY-BASED ONLINE ALGORITHM

### 5.1 Optimizing Power Based on a Probability Distribution

Let us assume that the length of the idle interval is generated by a fixed, known distribution whose density function is  $\pi$ , and devise a method to optimize power management decisions. We first discuss systems with two states and then give our generalization to the multistate case. Let  $\beta$  be the start-up energy of the sleep state, and  $\alpha$  the power dissipation of the active state. Suppose that the online algorithm uses  $\tau$  as the threshold at which time it will transition from the active state to the sleep state if the system is still idle. In this case, the expected energy cost for the algorithm for a single idle period will be

$$\int_0^\tau \pi(t)(\alpha t) dt + \int_\tau^\infty \pi(t)[\alpha\tau + \beta] dt. \quad (1)$$

The best online algorithm will select a value for  $\tau$  that minimizes this expression. The offline algorithm that knows the actual length of an upcoming idle period will have an expected cost of

$$\int_0^{\beta/\alpha} \pi(t)(\alpha t) dt + \int_{\beta/\alpha}^\infty \pi(t)\beta dt. \quad (2)$$

It is known for the two-state case that the online algorithm can pick its threshold  $\tau$  so that the ratio of its expected cost to the expected cost of the optimal algorithm is at most  $e/(e-1)$  [Karlin et al. 1990]. That is, for any  $\pi$  and any  $\alpha$  and  $\beta$ ,

$$\frac{\min_\tau \left\{ \int_0^\tau \pi(t)(\alpha t) dt + \int_\tau^\infty \pi(t)[\alpha\tau + \beta] dt \right\}}{\int_0^{\beta/\alpha} \pi(t)(\alpha t) dt + \int_{\beta/\alpha}^\infty \pi(t)\beta dt} \leq \frac{e}{e-1}. \quad (3)$$

This is optimal, in that for any  $\alpha$  and  $\beta$  there is a distribution  $\pi$  such that this ratio is at least  $e/(e-1)$ .

Let us now consider the multistate case. As in the previous section, let  $t_j$  be the solution to the equation  $\alpha_j t + \beta_j = \alpha_{j-1} t + \beta_{j-1}$ .  $t_j$  is the time that LEA will transition from state  $j-1$  to state  $j$ . We assume here that we have thrown out all the lines that do not appear on the lower envelope at some point. This is equivalent to the assumption that  $t_1 < t_2 < \dots < t_{k-1} < t_k$ . For ease of notation, we define  $t_0$  to be 0 and  $t_{k+1}$  to be  $\infty$ . The cost (expected energy consumption) of the optimal offline algorithm is

$$\sum_{i=0}^k \int_{t_i}^{t_{i+1}} \pi(t)[\alpha_i t + \beta_i] dt. \quad (4)$$

Now to determine the online algorithm, we must determine  $k$  thresholds, where the threshold  $\tau_i$  is the time at which the online algorithm will transition from state  $i - 1$  to state  $i$ . In the spirit of the deterministic online algorithm for the multistate case, we let  $\tau_i$  be the same as the threshold, which would be chosen if  $i - 1$  and  $i$  were the only two states. We call this algorithm the probability-based lower envelope algorithm (PLEA). The proof of the following theorem appears in the Appendix.

**THEOREM 2.** *For any distribution, the expected cost of the probability-based lower envelope algorithm is within a factor of  $e/(e - 1)$  of the expected cost for the optimal offline algorithm.*

## 5.2 Learning the Probability Distribution

While we have proven the competitive bounds regardless of the chosen idle time probability distribution, the practical problem of finding  $\pi(t)$  to guide the PLEA algorithm remains. Our approach is to learn  $\pi(t)$  online. Accordingly, the algorithm that uses PLEA in conjunction with our scheme to learn the probability distribution is called the online probability-based algorithm (OPBA). It works as follows: a window size  $w$  is chosen in advance and is used throughout the execution of the algorithm. The algorithm keeps track of the last  $w$  idle period lengths and summarizes this information in a histogram. Periodically, the histogram is used to generate a new power management strategy.

The set of all possible idle period lengths  $(0, \infty)$  is partitioned into  $n$  intervals, where  $n$  is the number of bins in the histogram. Let  $r_i$  be the left endpoint of the  $i$ th interval. The  $i$ th bin has a counter that indicates the number of idle periods among that last  $w$  idle periods whose length fell in the range  $[r_i, r_{i+1})$ . The bins are numbered from 0 through  $n - 1$ ,  $r_0 = 0$ , and  $r_n = \infty$ .

The last  $w$  idle periods are held in a queue. When a new idle period is completed, the idle period at the head of the queue is deleted from the queue. If this idle period falls in bin  $i$ , then the counter for bin  $i$  is decremented. The new idle period is added to the tail of the queue. If this idle period length falls into bin  $j$ , the counter for bin  $j$  is incremented. Thus, the histogram always includes data for the last  $w$  idle periods. Our experimental results in Section 7 include a study experimenting with values for  $w$ .

The counter for bin  $i$  is denoted by  $c_i$ . The threshold for changing states is selected among  $n$  possibilities: that is,  $r_0, \dots, r_{n-1}$  (the lower end of each range). We estimate the distribution  $\pi$  by the distribution that generates an idle period of length  $r_i$  with probability  $c_i/w$  for each  $i \in \{0, \dots, n - 1\}$ . The sum of the counters is the window length  $w$ . Thus, the threshold is taken to be

$$\arg \min_{r_i} \sum_{j=1}^{t-1} \left(\frac{c_j}{w}\right) r_j (\alpha_i - \alpha_{i-1}) + \sum_{j=t}^n \left(\frac{c_j}{w}\right) [r_t (\alpha_i - \alpha_{i-1}) + (\beta_{i-1} - \beta_i)]. \quad (5)$$

In order to be useful as a decision procedure, this algorithm must be implemented efficiently. We have implemented the algorithm for finding the all  $k$  thresholds in time  $O(kn)$ , where  $k + 1$  is the number of states and  $n$  is the

Table II. A Snapshot of the Histogram Used in OPBA. Note that the Offline Thresholds are 56, 2179, and 15 875, ms. The Number of Bins per State is 5

Bin	Range: Low End	Range: High End	Range Size	Count
1	0	11.2	11.2	35
2	11.2	22.4	11.2	2
3	22.4	33.6	11.2	4
4	33.6	44.8	11.2	7
5	44.8	<b>56</b>	11.2	4
6	56	478.8	422.8	5
7	478.8	901.6	422.8	3
8	901.6	1324.4	422.8	2
9	1324.4	1747.2	422.8	0
10	1747.2	<b>2170</b>	422.8	4
11	2170	4911	2741	7
12	4911	7652	2741	9
13	7652	10 393	2741	2
14	10 393	13 134	2741	5
15	13 134	<b>15 875</b>	2741	4
16	15 875	19 050	3175	2
17	19 050	22 225	3175	3
18	22 225	25 400	3175	1
19	25 400	28 575	3175	0
20	28 575	$\infty$	$\infty$	1

number of bins in the histogram. Two important factors determine the cost (in time expenditure) of implementing our method: (a) the frequency with which the thresholds are updated and (b) the number of bins in the histogram. These need to be minimized for policy implementation efficiency. The algorithm runs the  $O(kn)$  time algorithm every time the thresholds are updated. The frequency with which the thresholds are updated is the subject of one set of experiments discussed in Section 7. Note that the cost to implement our strategy is independent of  $w$ , the number of idle periods tracked in the histogram.

Minimization of the number of bins used in the histogram must be balanced with the fact that the finer grained the histogram, the more accurate our choice of thresholds will be. Our experiments, as discussed in the following sections, show that a finer grained binning is more important in some ranges than it is in others. A way of addressing this problem (key to the success of the algorithm) is to use the thresholds of the lower envelope algorithm to guide the selection of the bins. Recall that these thresholds are the  $t_1, \dots, t_k$  defined in Section 4. We then choose a constant  $c$  number of bins per state. The range from  $t_i$  to  $t_{i+1}$  is divided into  $c$  equal-sized bins. Note that the running time of our algorithm depends linearly on the number of bins, and hence depends linearly on  $c$ . For this reason, it is important to select a small value for  $c$ . We experimented with values for  $c$  ranging from 1 to 10 and the results only varied by 1%. We use a value of  $c = 5$  in all our results. Table II shows a sample histogram from our experiments.

Table III. This Figure Shows, for Each Trace, the Percentage of Idle Periods for Which the Optimal Algorithm Chose to Transition to Each State

Trace Length	State 1	State 2	State 3	State 4
34 575	0.804	0.158	0.011	0.024
68 139	0.673	0.233	0.092	2.9E-05
36 161	0.824	0.065	0.084	0.025
7250	0.727	0.045	0.057	0.169
10 648	0.787	0.096	0.061	0.054
7154	0.571	0.043	0.179	0.205
72 026	0.783	0.145	0.056	0.013
5130	0.643	0.155	0.063	0.137
46 929	0.78	0.152	0.031	0.034
24 821	0.59	0.208	0.151	0.048
9587	0.741	0.126	0.028	0.104
16 110	0.608	0.107	0.199	0.084
18 131	0.79	0.15	0.016	0.042
14 774	0.858	0.056	0.019	0.066

## 6. EXPERIMENTAL DESIGN

### 6.1 Data Used in Our Experiments

To demonstrate the utility of our probability-based algorithm, we use a mobile hard-drive from IBM [1996]. This drive has four power down states, as shown in Table I. Here, the start-up energy refers to the energy cost in transitioning from a state to the active state. For application disk access data, we used trace data from auspex file server archive.<sup>1</sup> From this data, we collected the arrival times and lengths for requests for disk access for 0.4 million disk accesses divided into multiple trace files, corresponding to different hours of the day.

Table III gives some data on these traces. The first column gives the number of requests. The subsequent columns give information about the behavior of the optimal algorithm when run on each trace. Specifically, they show for each state the percentage of idle periods in which the optimal algorithm transitions to that state. In all the traces, there is a high percentage of short sequences for which the optimal strategy is to stay in the active state (shown in column 2). The remaining percentages vary somewhat from trace to trace. All of the results reported in this paper are an average of the results on each individual trace, weighted by length.

### 6.2 Algorithm Test Suite

We compare OPBA and LEA to several other predictive algorithms presented in the literature. The algorithms come in two groups. The algorithms in the first group use a series of thresholds that determine when the algorithm will transition from each state to the next lower-power consumption state. OPBA and LEA fall into this group. The second group is made up of single-valued prediction

<sup>1</sup>Auspex File Traces from the NOW project. Available at. <http://now.cs.berkeley.edu/Xfs/AuspexTraces/auspex.html>.

Table IV. Predictions versus Outcomes for the Exponential Decay Algorithm. The Total Number of Idle Periods is 54 985 and the Total Number of Requests is 373 617

	Active	Stand-by	Idle	Sleep
Active	7631	3353	831	109
Stand-by	901	3144	4559	903
Idle	106	1414	1640	3416
Sleep	29	123	4897	7168

algorithms. They use a single prediction for the length of the upcoming idle period and to transition immediately to the optimal state for that length. They differ only in how they select a prediction for the length of the next idle period.

*Optimal Offline Algorithm (OPT)*: This algorithm is assumed to know the length of the idle period in advance. It selects the optimal power usage state for that idle period and then transitions to the active state before the new request arrives in order to service the incoming request just as it arrives.

*Last Period (LAST)*: This is a single-valued prediction algorithm that uses the last period as a predictor for the next idle period.

*Exponential Decay (EXP)*: This algorithm, developed by Hwang et al. [1996], keeps a single prediction for the upcoming idle period. After a new idle period ends, the prediction is updated by taking a weighted average of the old prediction and the new idle period length. Let  $p$  be the current prediction and  $l$  the length of the last idle period.  $p$  is updated as follows:

$$p \leftarrow \lambda p + (1 - \lambda)l,$$

where  $\lambda$  is a value in the range  $(0, 1)$ . We use a value of 0.5 for  $\lambda$ .

*Adaptive Learning Tree (TREE)*: This method uses an adaptive learning tree to predict the value of the next period based on the sequence of recent idle period lengths just observed. Details of this method can be found in Chung et al. [1999b].

There are different possible versions of single-valued prediction algorithms which are worth mentioning here. In describing these variations, we refer to the offline thresholds,  $t_1, \dots, t_k$ , described in Section 5.2. The authors of the learning tree algorithm observe that there are many idle periods that are very short. In order to avoid transitioning to a lower-powered state for such short idle periods, they keep their system in the active state until the first offline threshold  $t_1$  has been passed. Only then do they transition to the predicted optimal state.

We ran all single-valued prediction algorithms with and without this initial delay. We found that the results were not significantly different between the two versions. In this paper, we only report results for the versions without this delay. To see why the delay does not help significantly, we give some statistics for the exponential decay algorithm when run without the delay in Table IV. An entry in row  $i$  column  $j$  indicates the total number of times over all traces that the online algorithm predicted that the optimal state for an upcoming idle period would be  $j$  and the actual optimal state was  $i$ . Having a delay until time

$t_1$  helps only when the algorithm predicts that the optimal state is Stand-by, Idle, or Sleep and the actual optimal state is active. The number of times this happens is the sum of the values in row “active,” columns “Stand-by,” “Idle,” and “Sleep.” The degree of savings in latency is the largest for idle periods counted in entry [Active, Sleep], less so for entry [Active, Stand-by], and even less for [Active, Idle]. Given the values in the table, it is clear why having a delay does not offer a significant saving in the total latency.

Another feature that the authors of the learning tree algorithm employ is to transition to the active state after the threshold for the predicted optimal state is reached. This is done in a hope that a job will arrive shortly thereafter and the system can avoid incurring any additional latency in powering up after the new job has arrived. If no job arrives before the last threshold  $t_k$ , then the algorithm transitions to the lowest sleep state. To summarize, if there are  $k + 1$  states and the prediction is that state  $i$  will be the optimal state for the upcoming idle period, the algorithm will transition immediately to state  $i$ . If a new request has not arrived by time  $t_i$ , the algorithm will transition back to the active state (state 0). Finally, if a request has not arrived by time  $t_k$ , the algorithm will transition to the deepest sleep state, state  $k$ . We call this version of single-valued prediction algorithms the *preemptive-wake-up* version.

The alternative preemptive wake-up is to transition immediately to state  $i$  if that is the predicted optimal state and then to transition directly to the deepest sleep state if a request has not arrived by time  $t_i$ . We call this version the *non-preemptive-wake-up* version. Naturally, the preemptive-wake-up version will use more power, but will tend to incur less latency on average. We report results for both versions of all the single-valued predictive algorithms used in the study.

## 7. EXPERIMENTAL RESULTS

### 7.1 Experimentation with Window Size

Figure 2 shows the average energy consumed per request as the window size is varied. Note that beyond a certain threshold window size (below which the predictions are not accurate anyway) the variation of energy consumption with respect to the increase in window size is not large, indicating that our method is fairly robust to choices in window size. Our method performs best with a relatively small value for the window size, indicating that it is the most recent history that is the most relevant predicting upcoming idle period length. It also indicates that the distribution over idle period lengths is not necessarily stable over time. However, the results get worse if the window size gets below 10, showing that there need to be enough values to get a representative sample. We selected a value of 50 for the window size that is used for the remainder of the results presented.

### 7.2 Experimentation with Threshold Update Frequency

Figure 3 shows the average energy consumed per idle period as the frequency of updating the thresholds is varied. As one would expect, as the interval between

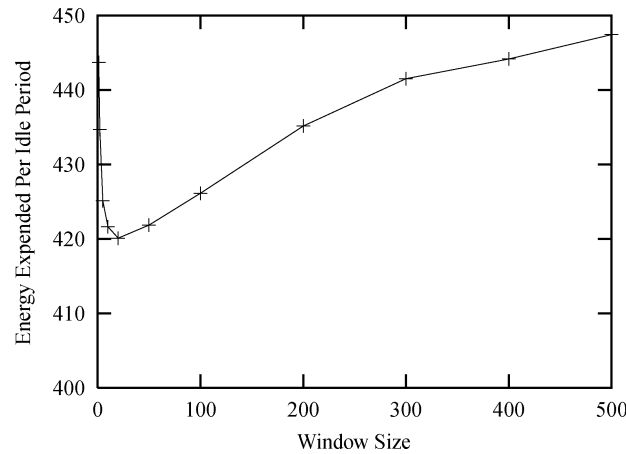


Fig. 2. Average energy consumed per request as a function of window size for the online probability-based algorithm. The thresholds are updated every ten requests.

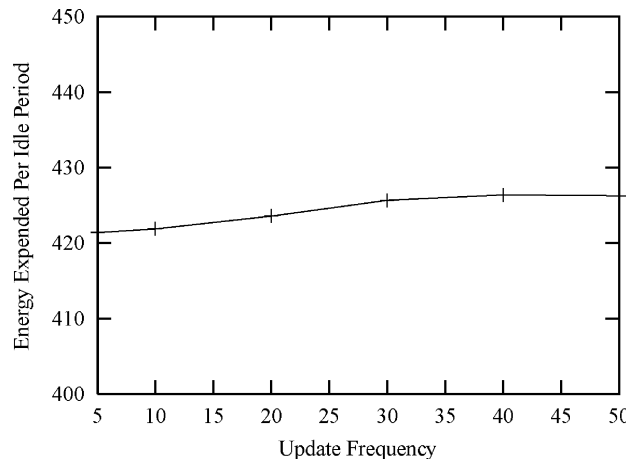


Fig. 3. Average energy consumed per request as a function of frequency of update for the online probability-based algorithm. The window size is 50.

updates grows, so does the power usage. However, there do not seem to be large differences in the cost, so we adopt a frequency of update of 50.

### 7.3 Evaluation of Algorithm Performance

Table V shows the comparison of energy consumption and wake-up latency effects across a number of predictive DPM algorithms. The first column of numbers in Table V is the average energy used per request for each of the algorithms. The second column is the ratio of this figure to the average energy consumed per request by the optimal offline algorithm. Interestingly, there were some traces where this ratio is less than 1 (although they always averaged out to be greater than 1 over all the traces). The reason it is possible for an algorithm to be better than the optimal offline algorithm on power consumption is because

Table V. Energy Measured in Joules and Latency Measured in Milliseconds

Algorithm	Average Energy	Competitive Ratio	Average Latency
Optimal	438.47	1	0
LEA	647.41	1.47	870.91
OPBA	446.45	1.01	1332.41
LAST: Preempt	868.51	1.9	1010.87
LAST: Non-preempt	477.96	1.09	2183.92
TREE: Preempt	990.53	2.25	1285.01
TREE: Non-preempt	460.44	1.05	2239.57
EXP: Preempt	550.75	1.25	1080.16
EXP: Non-preempt	458.98	1.04	1897.29

the optimal algorithm is always a forced wake-up preemptively before a request arrives. This means that the optimal algorithm incurs no additional latency due to waking up the disk drive. Recall that since there is a power-latency trade-off, this will tend to penalize the optimal algorithm with respect to power usage.

The average latency per request is shown in the final column of the figure. The latency is averaged over the number of requests, not the number of idle periods. These two numbers are different, since some requests may arrive while the device is busy working on other jobs or powering-up. In these cases, the incoming request would not correspond to an idle period. As an example, we refer back to the statistics for the exponential delay algorithm in Table IV. The total number of idle periods over all traces is 54 985, whereas the total number of requests is 373 617. Note that the algorithm only experiences the full 5 s of transition time when a request arrives and the algorithm is in the “Sleep” state. (Refer to Table I for the transition times for the various states.) For single-value prediction algorithms without predictive wake-up, this only happens if the algorithm predicts that the “Sleep” state will be the optimal state for the upcoming idle period or when the idle period actually lasts so long that the algorithm discovers that the sleep state was in fact the optimal state. The number of times this happens is the sum of the numbers in entries which are either in column “Sleep” or row “Sleep” (or both). The algorithm only experiences the 1.5 s delay in transitioning from the “Stand-by” state to the “Active” state for those idle periods counted in entries [Active, Stand-by], [Idle, Stand-by], and [Stand-by, Stand-by].

The final results in Table V are also shown graphically in Figure 4, which plots the power usage (middle column from Table V) against the average latency (last column from Table V). The OPBA exhibits the lowest power consumption among all the online algorithms. The other algorithms that come close to matching its performance in power (the nonpreemptive versions of LAST, TREE, and EXP) all suffer at least an additional 40% latency on average. Meanwhile, the algorithms that have a lower average latency than OPBA (LEA and the preemptive versions of LAST, TREE, and EXP) all use at least 25% more power on average. Thus, OPBA is the most successful algorithm in balancing power usage as well as latency incurred.

Finally, we study the the effect of variable wake-up time, since many devices vary somewhat in the time it takes to transition from one state to another.



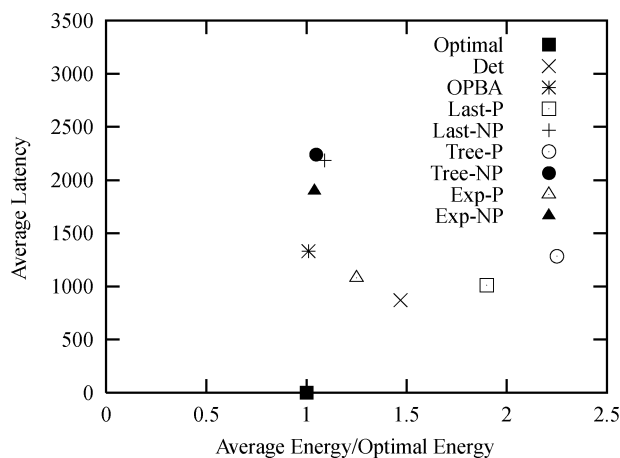


Fig. 4. Energy is measured in joules and latency is measured in milliseconds.

Table VI. Latency as Variation in Transition Time Increases. Each Transition Time is Multiplied by a Number Generated Uniformly at Random from the Range  $[1 - \epsilon, 1 + \epsilon]$ . Latency is measured in Milliseconds

$\epsilon$	LEA	OPBA	EXP-Preempt	EXP-Non-Preempt
0	879	1490	1092	1905
0.25	887	1543	1089	1945
0.5	894	1585	1108	2001
0.75	931	1631	1151	2116

We study the top performers for latency and power when random noise is added to the time to transition to the active state. Every time a transition is performed, we multiply the time to transition by a number generated uniformly at random from the range  $[1 - \epsilon, 1 + \epsilon]$  for different values of  $\epsilon$ . The results are shown in Table VI. We did not include the power usage because it does not vary significantly for different values of  $\epsilon$ . All of the methods suffer a slight degradation in performance as  $\epsilon$  increases. The variation seems to have a similar effect on all algorithms studied.

## 8. CONCLUSION AND FUTURE DIRECTIONS

We have presented a deterministic online algorithm for dynamic power management on multistate devices and proved that it is 2-competitive, and that this bound is tight. We improve upon this bound considerably by devising a probability-based scheme. To support this strategy in an online DPM framework, we provide a method to efficiently construct a probabilistic model for the length of the upcoming idle period based on online observations. Our experiments show that the algorithm presented in this paper attains the best performance in practice, in comparison to other known predictive DPM algorithm. The other algorithms that come close to matching its performance in power all suffer at least an additional 40% wake-up latency on average. Meanwhile, the

algorithms that have comparable latency to our methods all use at least 25% more power on average. Our future plans include extension of this work into systems with multiple active as well as power-down states.

Our experimental framework is available on our website<sup>2</sup> for testing a range of online DPM algorithms using a Java applet interface. The interface allows users to upload their data and evaluate our algorithms and other known algorithms that we have implemented in our simulation framework.

## APPENDIX

**PROOF OF THEOREM 1.** First, we establish that the worst case for the algorithm will always be just after a transition time. Consider the time  $t_j + \gamma$ , for some  $0 \leq j \leq k$  and  $0 \leq \gamma < t_{j+1} - t_j$ . For any value of  $\gamma$  in the given range, the optimal cost will be

$$\alpha_j(t_j + \gamma) + \beta_j.$$

For any value of  $\gamma$  in the given range, the online cost will be

$$\sum_{l=0}^{j-1} \alpha_l(t_{l+1} - t_l) + \alpha_j \gamma + \beta_j.$$

The ratio of these two will be maximized for  $\gamma = 0$ .

Now suppose that the interval ends just after  $t_j$  for some  $1 \leq j \leq k$ . Using the cost for the online and offline determined above, the ratio of the online cost to the offline cost will be

$$\frac{\sum_{l=0}^{j-1} \alpha_l(t_{l+1} - t_l) + \beta_j}{\alpha_j t_j + \beta_j} = 1 + \frac{\sum_{l=0}^{j-1} \alpha_l(t_{l+1} - t_l) - \alpha_j t_j}{\alpha_j t_j + \beta_j}.$$

Thus, it is sufficient to prove that

$$\sum_{l=0}^{j-1} \alpha_l(t_{l+1} - t_l) - \alpha_j t_j \leq \alpha_j t_j + \beta_j. \quad (6)$$

Each  $t_l$  was chosen so that

$$(\alpha_{l-1} - \alpha_l)t_l = \beta_l - \beta_{l-1},$$

so we can substitute these values into inequality (6) to get that

$$(\beta_1 - \beta_0) + (\beta_2 - \beta_1) + \cdots + (\beta_j - \beta_{j-1}) - \alpha_0 t_0 \leq \alpha_j t_j + \beta_j.$$

Collapsing the telescoping sum, we get that

$$\beta_j - \beta_0 - \alpha_0 t_0 \leq \alpha_j t_j + \beta_j.$$

<sup>2</sup>OSDPM Website at UC Irvine, <http://www.ics.uci.edu/osdpm>. "Java Applet based DPM Strategy Evaluation Website."

Using the fact that  $\beta_0 = 0$ ,  $t_0 = 0$ , and that  $\alpha_j \geq 0$  and  $t_j \geq 0$ , the inequality holds.  $\square$

**PROOF OF THEOREM 2.** Consider a system in which there are only two states:  $i - 1$  and  $i$ . Both online and offline must pay at least  $\alpha_i t$  for an interval of length  $t$ . In addition, each must pay at least  $\beta_{i-1}$  for the start-up cost. These costs, which are incurred by both algorithms regardless of their choices, will only serve to decrease the competitive ratio. In determining,  $\tau_i$ , we disregard these additional costs. Consider the system where the power consumption rate in the ON state is  $\alpha_{i-1} - \alpha_i$  and is 0 in the OFF state. The energy required to transition from the ON to the OFF state is  $\beta_i - \beta_{i-1}$ . We choose  $\tau_i$  to be the transition time for the optimal online policy in this system. Thus, we choose  $\tau_i$  to be

$$\arg \min_{\tau} \int_0^{\tau} \pi(t) t (\alpha_{i-1} - \alpha_i) dt + \int_{\tau}^{\infty} \pi(t) [\tau (\alpha_{i-1} - \alpha_i) + (\beta_i - \beta_{i-1})] dt.$$

The online cost for this new system is the above expression evaluated at  $\tau = \tau_i$ :

$$\text{ON}_i = \int_0^{\tau_i} \pi(t) t (\alpha_{i-1} - \alpha_i) dt + \int_{\tau_i}^{\infty} \pi(t) [\tau_i (\alpha_{i-1} - \alpha_i) + (\beta_i - \beta_{i-1})] dt.$$

Let  $t_i$  be defined to be  $(\beta_i - \beta_{i-1}) / (\alpha_{i-1} - \alpha_i)$ . Note that this is the same definition in the previous proof: the point where the lines  $\alpha_i t + \beta_i$  and  $\alpha_{i-1} t + \beta_{i-1}$  meet. The offline cost for the new system is

$$\text{OFF}_i = \int_0^{t_i} \pi(t) t (\alpha_{i-1} - \alpha_i) dt + \int_{t_i}^{\infty} \pi(t) (\beta_i - \beta_{i-1}) dt.$$

We are guaranteed that the ratio of the expected online to offline costs is at most  $e/(e - 1)$  [Karlin et al. 1990; Keshav et al. 1995].

Since the ratio of  $\text{ON}_i$  to  $\text{OFF}_i$  is at most  $e/(e - 1)$  for each  $i$ , we know that

$$\frac{\sum_{i=1}^k \text{ON}_i}{\sum_{i=1}^k \text{OFF}_i} \leq \frac{e}{e - 1}.$$

We now prove that  $\sum_{i=1}^k \text{ON}_i$  is exactly the expected cost for PLEA on the multilevel system. We also prove that  $\sum_{i=1}^k \text{OFF}_i$  is exactly the expected cost of the optimal algorithm for the multilevel system.

We rephrase  $\text{ON}_i$  by separating the integral into the intervals from  $\tau_{j-1}$  to  $\tau_j$ . To simplify notation,  $\tau_0$  denotes 0 and  $\tau_k$  denotes  $\infty$ .

$$\begin{aligned} \text{ON}_i &= \sum_{j=1}^i \int_{\tau_{j-1}}^{\tau_j} \pi(t) [t (\alpha_{i-1} - \alpha_i)] dt \\ &\quad + \sum_{j=i+1}^{k+1} \int_{\tau_{j-1}}^{\tau_j} \pi(t) [\tau_j (\alpha_{i-1} - \alpha_i) + (\beta_i - \beta_{i-1})] dt. \end{aligned}$$

In the sum over all  $\text{ON}_i$ , we group together all the contributions from each  $\text{ON}_i$  over the interval  $[\tau_j, \tau_{j+1}]$  for  $1 \leq j \leq k+1$ . Note that this is the interval that the algorithm will spend in state  $j$ . This value is

$$\sum_{i=1}^j \int_{\tau_j}^{\tau_{j+1}} \pi(t)[\tau_i(\alpha_{i-1} - \alpha_i) + (\beta_i - \beta_{i-1})] dt + \sum_{i=j+1}^k \int_{\tau_j}^{\tau_{j+1}} \pi(t)[t(\alpha_{i-1} - \alpha_i)] dt. \quad (7)$$

Thus, we have that

$$\sum_{i=1}^k \text{ON}_i = \sum_{j=0}^k f(j),$$

where

$$\begin{aligned} f(j) &= \sum_{i=1}^j \int_{\tau_j}^{\tau_{j+1}} \pi(t)[\tau_i(\alpha_{i-1} - \alpha_i) + (\beta_i - \beta_{i-1})] dt \\ &\quad + \sum_{i=j+1}^k \int_{\tau_j}^{\tau_{j+1}} \pi(t)[t(\alpha_{i-1} - \alpha_i)] dt. \end{aligned}$$

Putting the summations inside the integrals and collapsing the telescoping sums, the expression in (7) becomes

$$\int_{\tau_j}^{\tau_{j+1}} \pi(t) \text{cost}(t) dt,$$

where

$$\text{cost}(t) = (\beta_j - \beta_0) + \tau_1 \alpha_0 + \sum_{l=1}^{j-1} (\tau_{l+1} - \tau_l) \alpha_l + (t - \tau_l) \alpha_l.$$

Note that

$$(\beta_j - \beta_0) + \tau_1 \alpha_0 + \sum_{l=1}^{j-1} (\tau_{l+1} - \tau_l) \alpha_l + (t - \tau_l) \alpha_l$$

is exactly the energy expended by PLEA if the idle period  $t$  is in the range  $[\tau_j, \tau_{j+1}]$ . Thus, the expected cost for PLEA is

$$\sum_{j=0}^k \int_{\tau_j}^{\tau_{j+1}} \pi(t) \text{cost}(t) dt = \sum_{i=1}^k \text{ON}_i.$$

The proof that the expected offline cost is equal to  $\sum_{i=1}^k \text{OFF}_i$  is the same as the proof for the online cost, except that the integrals are separated into intervals according to the  $t_i$  instead of the  $\tau_i$ .

## REFERENCES

- BENINI, L., BOGLIOLO, A., AND DE MICHELI, G. 2000. A survey of design techniques for system-level dynamic power management. *IEEE Transactions on Very Large Scale Integration (TVLSI) Systems* 8, 3, 299–316.

- BENINI, L., BOGLIOLO, A., PALEOLOGO, G., AND DE MICHELI, G. 1999. Policy optimization for dynamic power management. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 18, 6, 813–833.
- BENINI, L. AND DE MICHELI, G. 1998. *Dynamic Power Management: Design Techniques and CAD Tools*. Kluwer Dordrecht.
- BENINI, L., DE MICHELI, G., AND MACII, E. 2001. Designing low-power circuits: Practical recipes. *IEEE Circuits and Systems Magazine* 1, 1 (March), 6–25.
- BORODIN, A. AND EL-YANIV, R. 1998. *Online Computation and Competitive Analysis*. Cambridge University Press, Cambridge, MA.
- CHUNG, E. Y., BENINI, L., BOGLIOLO, A., AND DE MICHELI, G. 1999a. Dynamic power management for non-stationary service requests. In *Proceedings of the Design Automation and Test Europe*.
- CHUNG, E.-Y., BENINI, L., AND DE MICHELI, G. 1999b. Dynamic power management using adaptive learning trees. In *Proceedings of ICCAD*.
- GLYNN, P., SIMUNIC, T., BENINI, L., AND DE MICHELI, G. 2000. Dynamic power management of portable systems. In *Proceedings of MOBICOM*.
- HWANG, C.-H., ALLEN, C., AND WU, H. 1996. A predictive system shutdown method for energy saving of event-driven computation. In *Proceedings of the IEEE/ACM International Conference on Computer Aided Design*. 28–32.
- IBM. 1996. *Technical Specifications of Hard Drive IBM Travelstar VP 2.5 inch*. Available at, <http://www.storage.ibm.com/storage/oem/data/travvp.htm>.
- IRANI, S., SHUKLA, S., AND GUPTA, R. 2002. Competitive analysis of dynamic power management strategies for systems with multiple power saving states. In *Proceedings of the Design Automation and Test Europe Conference*.
- KARLIN, A., MANASSE, M., MCGEOCH, L., AND OWICKI, S. 1990. Randomized competitive algorithms for non-uniform problems. In *First Annual ACM-SIAM Symposium on Discrete Algorithms*. 301–309.
- KESHAV, S., LUND, C., PHILLIPS, S., REAINGOLD, N., AND SARAN, H. 1995. An empirical evaluation of virtual circuit holding time policies in ip-over-atm networks. *IEEE Journal on Selected Areas in Communications* 13, 1371–1382.
- LU, Y., CHUNG, E. Y., SIMUNIC, T., BENINI, L., AND DE MICHELI, G. 2000. Quantitative comparison of power management algorithms. In *DATE—Proceedings of the Design and Automation and Test in Europe Conference and Exhibition*.
- LU, Y. AND DE MICHELI, G. 1999. Adaptive hard disk power management on personal computers. In *Proceedings of the Great Lakes Symposium on VLSI*.
- PHILLIPS, S. J. AND WESTBROOK, J. R. 1999. On-line algorithms: Competitive analysis and beyond. In *Algorithms and Theory of Computation Handbook*. CRC Press, Boca Raton, FL.
- QIU, Q. AND PEDRAM, M. 1999. Dynamic power management based on continuous-time Markov decision processes. In *Proceedings of Design Automation Conference (June)* 555–561.
- QIU, Q., WU, Q., AND PEDRAM, M. 1999. Stochastic modeling of a power-managed system: Construction and optimization. In *Proceedings of the International Symposium on Low Power Electronics and Design*.
- QIU, Q., WU, Q., AND PEDRAM, M. 2000. Dynamic power management of complex systems using generalized stochastic petri nets. In *Proceedings 37th Design Automation Conference (DA'2000)* (5–9 June). Los Angeles, CA. 352–356.
- RAMANATHAN, D. 2000. *High-Level Timing and Power Analysis for Embedded Systems*. In PhD thesis, University of California at Irvine.
- RAMANATHAN, D., IRANI, S., AND GUPTA, R. K. 2000. Latency effects of system level power management algorithms. In *Proceedings of the IEEE International Conference on Computer-Aided Design*.
- RAMANATHAN, D., IRANI, S., AND GUPTA, R. 2002. An analysis of system level power management algorithms and their effects on latency. *IEEE Transactions on Computer Aided Design* 21, (March) 3.
- SHUKLA, S. AND GUPTA, R. 2001. A model checking approach to evaluating system level power management for embedded systems. In *Proceedings of IEEE Workshop on High Level Design Validation and Test (HLDVT01)* (November). IEEE Press, New York.

SIMUNIC, T., BENINI, L., AND DE MICHELI, G. 1999. Event driven power management of portable systems. In *Proceedings of International Symposium on System Synthesis*. 18–23.

SRIVASTAVA, M. B., CHANDRAKASAN, A. P., AND BRODERSON, R. W. 1996. Predictive shutdown and other architectural techniques for energy efficient programmable computation. *IEEE Transactions on VLSI Systems* 4, 1 (March), 42–54.

Received February 2002; accepted July 2002