GUIDELINES FOR DATA-PARALLEL CYCLE-STEALING IN NETWORKS OF WORKSTATIONS, I

A.L. ROSENBERG

CMPSCI TR 98-15

March 1998

Guidelines for Data-Parallel Cycle-Stealing in Networks of Workstations, I*

Arnold L. Rosenberg
Department of Computer Science
University of Massachusetts
Amherst, MA 01003
rsnbrg@cs.umass.edu

Abstract. We derive guidelines for nearly optimal scheduling of data-parallel computations within a draconian mode of cycle-stealing in NOWs. In this computing regimen, workstation A takes control of workstation B's processor whenever it is idle, with the promise of relinquishing control *immediately* upon demand—thereby losing work in progress. The typically high communication overhead for supplying workstation B with work and receiving its results militates in favor of supplying B with large amounts of work at a time; the risk of losing work in progress when B is reclaimed militates in favor of supplying B with a succession of small bundles of work. The challenge is to balance these two pressures in a way that maximizes (some measure of) the amount of work accomplished. Our guidelines attempt to maximize the expected work accomplished by workstation B in an episode of cycle-stealing, assuming knowledge of the instantaneous probability of workstation B's being reclaimed. Our study is a step toward rendering prescriptive the descriptive study of cycle-stealing in [3].

1 The Cycle-Stealing Problem

We derive guidelines for (almost) optimally scheduling data-parallel computations on "borrowed" workstations, within the model developed in [3]. The phenomenological study in that paper builds on the following rather draconian version of cycle-stealing in networks of workstations (NOWs)—the use by one workstation of idle computing cycles of another. The owner of workstation A contracts to take control of workstation B whenever

* A portion of this work was presented at the *12th Intl. Parallel Proc. Symp.*, March 30 - April 3, 1998, Orlando, FL.

its owner is absent. When the owner of B reclaims that workstation, workstation A *immediately* relinquishes control of B, killing any active job(s)—thereby destroying all work since the last checkpoint.

> Such draconian "contracts" are inevitable, for instance, when a returning owner unplugs a laptop from a network; one encounters such contracts also at several institutions where cycle-stealing is supported.

Such a "contract" creates a tension between the following inherently conflicting aspects of cycle-stealing. On the one hand, since any work in progress on workstation B when it is reclaimed is lost, a cycle-stealer wants to break a cycle-stealing episode into many short *periods*, supplying small amounts of work to the borrowed workstation each time. On the other hand, since each of the inter-workstation communications that bracket every period in a cycle-stealing episode—to supply work to workstation B and to reclaim the results of that work—involves an expensive setup protocol, the cycle-stealer wants to break each cycle-stealing episode into a few long periods, supplying large amounts of work to workstation B each time. Clearly, the challenge in scheduling episodes of cycle-stealing is to balance these conflicting factors in a way that maximizes the productive output of the episode. The research we report on here resolves the preceding conflict by deriving scheduling guidelines that (approximately) maximize the expected work[1] accomplished within an episode of cycle-stealing, within the following setting. We focus on computations that are data-parallel, in that they consist of a massive number of independent repetitive tasks of known durations.

> One encounters such computations in many scientific applications.

We develop our schedules under the assumption that we know the instantaneous probability of workstation B's being reclaimed and that the function yielding this information is "smooth."

> Although our results are stated as though we had exact knowledge of these probabilities, they extend easily to situations wherein this knowledge is approximate, garnered possibly from trace data that exposes B's owner's computer usage patterns. Our assumption about "smoothness" is reasonable, since one would likely encapsulate even trace data by some "well-behaved" curve.

[1]In a forthcoming sequel to this paper, we focus on (nearly) optimizing a worst-case, rather than expected, measure of a cycle-stealing episode's work output.

Our hope—and experience—is that the approximate specifications one obtains via the guidelines derived here provide one with a manageably narrow search space for a truly optimal schedule.

A roadmap. Section 2 presents the formal model under which we derive our scheduling guidelines. In Section 3 we present the results about the structure of optimal schedules that underlie the guidelines. In Section 4, we illustrate the application of the guidelines in a variety of scenarios, and we compare the resulting schedules with the (ad hoc) provably optimal ones from [3]. We present further results about the structure of optimal schedules in Section 5, both to lend insight and perspective to our guidelines and to supply raw material for certain ad hoc improvements to the schedules they prescribe. We end with suggestions for further research in Section 6.

Related work. Other noteworthy studies of scheduling algorithms for NOWs, which differ from ours in focus or objectives, appear in [1, 2, 4, 5, 6]. Of these, only [2] deals with the present adversarial scenario of stealing cycles; its main contribution is a randomized strategy that, with high probability, steals cycles within a logarithmic factor of optimally. We do not list the many empirical studies of computation on NOWs whose main foci are on enabling systems or specific applications rather than on analyzed scheduling algorithms.

Remark. The model we study here has applications to "real-life" problems other than scheduling single episodes of cycle-stealing. One important example is scheduling saves in a fault-prone computing system, as studied in [7]. This problem admits an abstract formulation that is formally similar to our model for cycle-stealing. Our model differs in many details from that of [7], and our research methodology differs dramatically from that paper's, but it is clear that our results can be adapted to apply in that setting also.

2 A Formal Model of Data-Parallel Cycle-Stealing

We review the basic structure of the cycle-stealing model of [3], focusing only on details that are relevant to the current study. We refer the reader to that paper for additional details and variations on the model presented here.

2.1 The Model

Overview. We schedule data-parallel cycle-stealing in an "architecture-independent" fashion, in the sense of [9]: the cost of inter-workstation communications is characterized by a single (overhead) parameter c, which is the (combined) cost of initiating both the communication in which workstation A sends work to workstation B, and the communication in which B returns the results of the work. We assume that: tasks are indivisible; task times may vary but are known perfectly; the time for a task includes the marginal cost of transmitting its input and output data (so we may keep c independent of the sizes of data transmissions).

Cycle-stealing schedules. Workstation A schedules an episode of cycle-stealing by partitioning the time of B's potential availability into a sequence of nonoverlapping *periods.* For simplicity, we identify a cycle-stealing schedule $\mathcal{S}$ with its sequence of period-lengths: $\mathcal{S} = t_0, t_1, \ldots$, where each $t_i > 0$. A schedule can be finite, when there is a known upper bound L on the length of the episode, or it can be infinite, when no such bound is known. (Examples of both situations appear later.) The intended interpretation is that at time

$$\tau_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } k = 0 \\ T_{k-1} \stackrel{\text{def}}{=} t_0 + t_1 + \cdots + t_{k-1} & \text{if } k > 0 \end{cases}$$

the kth period begins: workstation A supplies workstation B with an amount of work chosen so that t_k time units are sufficient for A to send the work to B, for B to perform the work, and for B to return the results of the work to A.

The work achieved by a schedule. The amount of work achievable in a period of length t_k is[2] $t_k \ominus c$. If workstation B *is not reclaimed* by time $T_k = \tau_k + t_k$, then the amount of work done so far during the episode is augmented by $t_k \ominus c$; if B *is reclaimed* by time T_k, then the episode ends, having accomplished work $\sum_{i=0}^{k-1}(t_i \ominus c)$.

> The limits of the latter summation implicitly reflect both the loss of work from the interrupted period and the termination of the episode.

Easily, the risk of having a period interrupted, thereby losing work, may make it desirable to have the lengths of a bounded-lifespan schedule's *productive* periods (those with $t_i > c$) sum to *less than* the potential duration L of the episode.

Cycle-stealing with known risk. One cannot derive provably productive scheduling guidelines without some antidote for a malicious adversary who kills every episode of cycle-stealing just at the end of the 0th period. Here (as in the first half of [3]), this antidote resides in our assumed knowledge of the risk of being interrupted in the midst of a period, in the form of the *life function* p of an episode: for each time t, $p(t)$ is the probability that workstation B *has not been reclaimed by time* t. In accord with the motivating scenario: $p(0) = 1$; when an upper bound L to the duration of the episode

[2] The operator "$\ominus$" denotes *positive subtraction* and is defined by: $x \ominus y \stackrel{\text{def}}{=} \max(0, x - y)$.

is known, then p decreases monotonically to 0 in the range[3] $[0, L]$; when no bound L is known, then p decreases monotonically for all t, with the limit 0. In order to enable our analytical results, we let period-lengths be arbitrary real numbers, and we consider only life functions that are "well-behaved," in the sense of being *differentiable* along the entire real axis and of *having no flex points*. These idealizations make even our "definitive" results just guidelines.

Our goal is to maximize the *expected work* in an episode of cycle-stealing. For any schedule $\mathcal{S} = t_0, t_1, \ldots$ and life function p, this quantity is given by

$$E(\mathcal{S}; p) \stackrel{\text{def}}{=} \sum_{i \geq 0} (t_i \ominus c) p(T_i). \tag{2.1}$$

The summation here has upper limit $m-1$ for an m-period schedule and ∞ for an infinite schedule (when no duration bound is known). The simple functional form of $E(\mathcal{S})$ makes it easy to cope with life functions that are known only approximately (say, via trace data).

Cycle-stealing schedule $\mathcal{S}^\star$ is *optimal for life function* p if it maximizes the expected production of work, $E(\mathcal{S}; p)$, over all schedules $\mathcal{S}$ for p. The guidelines we derive emerge from exposing the structure of optimal schedules.

2.2 A Useful Technical Result

The analysis that leads to our scheduling guidelines is simplified by the fact that we lose no generality by focusing only on schedules that are *productive*, in the sense of the schedule $\mathcal{S}'$ of the following result from [3]. We state and use a slightly stronger version of the result than one finds in [3]; the proof there actually supports this version. Importantly, this result permits us henceforth to use ordinary subtraction rather than positive subtraction in calculations involving expression (2.1).

Proposition 2.1 ([3]) *Any schedule $\mathcal{S}$ for a life function p can be replaced by a schedule $\mathcal{S}'$ such that:*

- $E(\mathcal{S}'; p) \geq E(\mathcal{S}; p)$;
- *each period of $\mathcal{S}'$—save the last, if $\mathcal{S}'$ is finite—has length $> c$.*

[3] As usual, the assertion "$a \in [b, c]$" (resp., "$a \in (b, c)$") is equivalent to "$b \leq a \leq c$" (resp., "$b < a < c$").

3 Our Scheduling Guidelines

We derive our scheduling guidelines by characterizing the dependencies among the period-lengths of optimal schedules for "smooth" (i.e., differentiable) life functions p. We first state the results that underlie the guidelines, in Section 3.1, and then prove the results in the Sections 3.2 and 3.3.

3.1 The Structure of Optimal Schedules

Our main results presuppose "nice" structure in the life function p; all require that p be differentiable; some require additionally that p enjoy one of the following nice "shapes."

> The life function p is *concave* (resp., *convex*) if its derivative is everywhere nonincreasing (resp., everywhere nondecreasing): for all positive real ξ and $\eta > \xi$, we have $p'(\xi) \geq p'(\eta)$ (resp., $p'(\xi) \leq p'(\eta)$).

The three life functions studied in [3] illustrate these properties. (1) The *geometrically increasing risk* life function with potential lifespan L, $p(t) \stackrel{\text{def}}{=} (2^L - 2^t)/(2^L - 1)$, is concave. The risk of interruption doubles at every time unit in this scenario. (2) The *geometrically decreasing lifespan* life function with risk factor $a > 1$, $p(t) \stackrel{\text{def}}{=} a^{-t}$, is convex. Each episode in this scenario has a "half-life." (3) The *uniform-risk* life function with potential lifespan L, $p(t) \stackrel{\text{def}}{=} 1 - t/L$, is both concave and convex. The risk of interruption in this scenario is uniform across the potential lifespan. We shall revisit these scenarios in Section 4.

Our main results. Say that the schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for the differentiable life function p.

1. A characterization of the optimal sequence of t_i's:

> *The period-lengths of $\mathcal{S}$ are given implicitly by the inductive system of equations: for each period-index $k \geq 0$,*
>
> $$p(T_k) \; = \; -\sum_{j \geq k} (t_j - c) p'(T_j).$$
>
> *In computationally more useful form: for each period-index $k > 0$,*
>
> $$p(T_k) \; = \; p(T_{k-1}) + (t_{k-1} - c) p'(T_{k-1}).$$

2. Bounds on the optimal t_0:

If the life function p is convex*, then*

$$\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0)}} + \frac{c}{2} \;<\; t_0 \;\leq\; 2\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0)}} + c.$$

If the life function p is concave*, then*

$$\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0)}} + \frac{c}{2} \;<\; t_0 \;\leq\; 2\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0/2)}} + c.$$

The next two subsections are devoted to proving these results.

3.2 The Dependencies among the Optimal Period-Lengths

Theorem 3.1 *If schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for the differentiable life function p, then the period-lengths of $\mathcal{S}$ are given implicitly by the following system of equations. For each period-index $k \geq 0$,*[4]

$$p(T_k) \;=\; -\sum_{j \geq k} (t_j - c)p'(T_j). \tag{3.1}$$

Proof. Theorem 3.1 follows from the fact that schedule $\mathcal{S}$, being optimal, accomplishes at least as much work as does any "shifted" version of $\mathcal{S}$. This claim is formalized and verified as follows.

> The $\langle k, -\delta\rangle$-*shift*, $\mathcal{S}^{\langle k,-\delta\rangle}$, of $\mathcal{S}$ and the $\langle k, +\delta\rangle$-*shift*, $\mathcal{S}^{\langle k,+\delta\rangle}$, of $\mathcal{S}$ are the schedules
>
> $$\mathcal{S}^{\langle k,-\delta\rangle} \stackrel{\text{def}}{=} t_0, t_1, \ldots, t_{k-1}, t_k - \delta, t_{k+1}, \ldots,$$
> $$\mathcal{S}^{\langle k,+\delta\rangle} \stackrel{\text{def}}{=} t_0, t_1, \ldots, t_{k-1}, t_k + \delta, t_{k+1}, \ldots,$$
>
> which have the same number of periods[5] as $\mathcal{S}$ and the same period-lengths, save for period k.

[4] The upper limit of each summation in (3.1) is inherited from summation (2.1).

[5] That is, if $\mathcal{S}$ has finitely many periods, then $\mathcal{S}^{\langle k,+\delta\rangle}$ and $\mathcal{S}^{\langle k,-\delta\rangle}$ have the same number; if $\mathcal{S}$ has infinitely many periods, then so also do $\mathcal{S}^{\langle k,+\delta\rangle}$ and $\mathcal{S}^{\langle k,-\delta\rangle}$.

(a) We first compare $E(\mathcal{S};p)$ with $E(\mathcal{S}^{\langle k,-\delta\rangle};p)$. By direct calculation and the assumed optimality of schedule $\mathcal{S}$, we find that

$$0 \ \leq \ E(\mathcal{S};p) - E(\mathcal{S}^{\langle k,-\delta\rangle};p) \ = \ \delta p(T_k - \delta) + \sum_{i\geq k} (t_i - c)\left[p(T_i) - p(T_i - \delta)\right],$$

so that

$$p(T_k - \delta) \ \geq \ -\sum_{i\geq k} (t_i - c)\left[\frac{p(T_i) - p(T_i - \delta)}{\delta}\right]. \tag{3.2}$$

Since an inequality of the form (3.2) holds for every period-index k and for arbitrarily small δ, the differentiability of p implies that[6]

$$p(T_k) \ \geq \ -\sum_{i\geq k} (t_i - c)p'(T_i). \tag{3.3}$$

(b) We next compare $E(\mathcal{S};p)$ with $E(\mathcal{S}^{\langle k,+\delta\rangle};p)$. Mirroring the reasoning in part (a), we derive the chain

$$0 \ \geq \ E(\mathcal{S}^{\langle k,+\delta\rangle};p) - E(\mathcal{S};p) \ = \ \delta p(T_k + \delta) + \sum_{i\geq k} (t_i - c)\left[p(T_i + \delta) - p(T_i)\right],$$

so that

$$p(T_k + \delta) \ \leq \ -\sum_{i\geq k} (t_i - c)\left[\frac{p(T_i + \delta) - p(T_i)}{\delta}\right]. \tag{3.4}$$

As before, since p is differentiable, the fact that inequality (3.4) holds for every period-index k and for arbitrarily small δ implies that

$$p(T_k) \ \leq \ -\sum_{i\geq k} (t_i - c)p'(T_i). \tag{3.5}$$

Inequalities (3.3) and (3.5) combine to yield the system (3.1). ■

The implicit specifications (3.1) of the period-lengths of $\mathcal{S}$ are often computationally difficult to instantiate for specific p, depending on the functional form of p. We therefore present, in the following corollary, more easily instantiated specifications for all period-lengths t_k of $\mathcal{S}$ save the first. We shall have to deal with the first period-length, t_0, separately, in Section 3.3.

Corollary 3.1 *If schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for the differentiable life function p, then, for each period-index $k \geq 1$,*

$$p(T_k) \ = \ p(T_{k-1}) + (t_{k-1} - c)p'(T_{k-1}). \tag{3.6}$$

[6]Note that we are implicitly letting δ tend to 0 here.

Proof. The system of equations (3.6) is easily derived from (3.1), via the following easily derived intermediate system of equations.

$$p(T_k) \;=\; p(T_0) + \sum_{j=0}^{k-1}(t_j - c)p'(T_j),$$

for each period-index $k \geq 1$ of $\mathcal{S}$. ■

We end this subsection with a purely technical result that bolsters the observation in [3] that not every life function admits an optimal schedule. Corollary 3.1 combines with Proposition 2.1 to yield a simple test for such schedules.

Corollary 3.2 *If the life function p admits an optimal schedule, then there exists $t > c$ such that $p(t) \;>\; -(t-c)p'(t)$.*

Using this result, one can show, for instance, that life functions of the form $p(t) = 1/(t+1)^d$, where $d > 1$, do not admit optimal schedules.

3.3 Nearly Tight Bounds on the Optimal Initial Period-Length

Whereas we were able to convert system (3.1) into the computationally friendly prescription (3.6) for inductively optimally selecting all period-lengths t_k save the initial one, we have been able only to derive upper and lower bounds on the optimal initial period-length t_0. This section is devoted to establishing these bounds. Importantly, these bounds are computationally more friendly than the $(k = 0)$-instance of system (3.1). Unfortunately, though, our two upper bounds are established only for life functions that enjoy additional structural uniformity (either convexity or concavity). As we indicate in Section 4, our bounds on t_0 are often moderately tight, hence combine with system (3.1) to yield useful guidelines for determining an optimal schedule for smooth life functions.

3.3.1 A Lower Bound on the Optimal t_0

Our lower bound on the optimal first period-length t_0 actually holds for general differentiable life functions.

Theorem 3.2 *If the schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for life function p, then*

$$t_0 \;\geq\; \sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0)}} + \frac{c}{2}. \tag{3.7}$$

Proof. If schedule $\mathcal{S}$ is optimal for p, then $E(\mathcal{S};p) \geq E(\widetilde{\mathcal{S}};p)$, where schedule $\widetilde{\mathcal{S}} \stackrel{\text{def}}{=} t_0 + t_1, t_2, \ldots$ is obtained from $\mathcal{S}$ by combining the first two periods. We thus have

$$0 \ \leq \ E(\mathcal{S};p) - E(\widetilde{\mathcal{S}};p) \ = \ (t_0 - c)p(t_0) - t_0 p(T_1). \tag{3.8}$$

If we now combine the rightmost expression in (3.8) with the $(k=0)$-instance of system (3.6), and we note the positivity of the result, then we find that

$$(t_0 - c)p(t_0) - t_0 p(T_1) \ = \ -cp(t_0) - t_0(t_0 - c)p'(t_0) \ \geq \ 0. \tag{3.9}$$

"Solving" (3.9) for t_0 (by completing the square) now yields (3.7). ■

3.3.2 Upper Bounds on the Optimal t_0

Our first upper bound on t_0 is an implicit one which holds for general differentiable life functions. Its main utility is in deriving explicit upper bounds for life functions having special "shapes."

Lemma 3.1 *Let $\mathcal{S} = t_0, t_1, \ldots$ be an optimal schedule for life function p. Either the initial period-length $t_0 \leq 2c$, or t_0 is small enough that*

$$p(t_0) \ \geq \ \max_{t \in (c,\ t_0 - c)} \left(1 - \frac{c}{t}\right) p(t). \tag{3.10}$$

Proof. Assume, for contradiction, that $t_0 > 2c$ and that t_0 is so large that condition (3.10) is violated; i.e., there is a $\hat{t} \in (c,\ t_0 - c)$ for which

$$p(t_0) < (1 - c/\hat{t})p(\hat{t}). \tag{3.11}$$

Define the schedule $\widehat{\mathcal{S}} \stackrel{\text{def}}{=} \hat{t}, t_0 - \hat{t}, t_1, \ldots$; that is, schedule $\widehat{\mathcal{S}}$ agrees with schedule $\mathcal{S}$, except that it splits $\mathcal{S}$'s initial period in two. By direct calculation, we have

$$\begin{aligned} E(\widehat{\mathcal{S}};p) - E(\mathcal{S};p) \ &= \ (\hat{t} - c)p(\hat{t}) + (t_0 - \hat{t} - c)p(t_0) - (t_0 - c)p(t_0) \\ &= \ (\hat{t} - c)p(\hat{t}) - \hat{t}p(t_0). \end{aligned} \tag{3.12}$$

By (3.11), the difference (3.12) is *strictly* positive. This contradicts the assumed optimality of $\mathcal{S}$, hence establishes (3.10). ■

Lemma 3.1 leads to the following explicit, computationally useful, upper bounds on the optimal t_0. These bounds, which require that the life function p be either convex or concave, combine with the lower bound (3.7) to bracket t_0 for many "smooth" life functions within a factor of 2.

Theorem 3.3 *Say that schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for the life function p and that $t_0 > 2c$. If p is* convex, *then*

$$t_0 \;\leq\; 2\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0)}} + c. \tag{3.13}$$

If p is concave, *then*

$$t_0 \;\leq\; 2\sqrt{\frac{c^2}{4} - \frac{cp(t_0)}{p'(t_0/2)}} + c. \tag{3.14}$$

Proof. If we instantiate the system of inequalities that is implicit in (3.10) with the value $t = \frac{1}{2}t_0$, then we find that

$$\left(1 - \frac{2c}{t_0}\right)[p(t_0) - p(t_0/2)] \;\geq\; -\frac{2c}{t_0}p(t_0).$$

Invoking the MVT on the lefthand side of this inequality, we find that there exists a $\xi \in (t_0/2,\ t_0)$ such that

$$t_0\left(\frac{1}{2}t_0 - c\right)p'(\xi) \;\geq\; -2cp(t_0). \tag{3.15}$$

Note now that, if the life function p is convex, then $p'(\xi) \leq p'(t_0)$, while, if the life function p is concave, then $p'(\xi) \leq p'(t_0/2)$. Hence, when p is convex, inequality (3.15) yields[7]

$$t_0\left(\frac{1}{2}t_0 - c\right) \;\leq\; -2c\frac{p(t_0)}{p'(t_0)}; \tag{3.16}$$

and, when p is concave, inequality (3.15) yields

$$t_0\left(\frac{1}{2}t_0 - c\right) \;\leq\; -2c\frac{p(t_0)}{p'(t_0/2)}. \tag{3.17}$$

If we now "solve" inequalities (3.16) and (3.17) for t_0 (by completing the square), then we derive inequality (3.13) for the case of convex p and inequality (3.14) for the case of concave p. ■

[7]The inequality reverses as we go from (3.15) to (3.16) and (3.17) because p' is negative.

4 Optimal Schedules vs. Guideline-Generated Ones

We illustrate the utility of the guidelines derived in Section 3 by applying them to some specific life functions, producing for each an approximation to the optimal schedule $\mathcal{S} = t_0, t_1, \ldots$. We focus mostly on life functions for which we know absolutely optimal schedules, via the study in [3]. Using system (3.6), we easily derive explicit expressions for each non-initial period-length t_k, where $k \geq 1$, in terms of all preceding period-lengths. Of course, these expressions are explicit only modulo our finding an explicit expression for t_0. We can only approximate this latter task, by using the lower bound (3.7), in conjunction with whichever of the upper bounds (3.13) and (3.14) is appropriate for the life function(s) in question.

4.1 The Family $\{p_{d,L}(t) \stackrel{\text{def}}{=} 1 - t^d/L^d \mid d = 1, 2, \ldots\}$

We begin studying a family of concave life functions for an episode of cycle-stealing with potential lifespan L; the $(d = 1)$-member of the family is the life function for the *uniform risk scenario* of [3], wherein the risk of interruption is stable across the opportunity.

The non-initial period-lengths. The kth period-length t_k of an optimal schedule $\mathcal{S} = t_0, t_1, \ldots$ for $p_{d,L}$ can be determined as follows. By system (3.6), we have

$$L^d - (T_{k-1} + t_k)^d \;=\; L^d - T_{k-1}^d - d(t_{k-1} - c)T_{k-1}^{d-1},$$

which simplifies to

$$t_k \;=\; \left(\left(1 + \frac{d(t_{k-1} - c)}{T_{k-1}} \right)^{1/d} - 1 \right) T_{k-1}.$$

When $d = 1$, this expression simplifies even further, to

$$t_k \;=\; t_{k-1} - c, \tag{4.1}$$

which is identical to the optimal period-length recurrence for $p_{1,L}$ discovered in [3].

The initial period-length. We now invoke inequality (3.7) to obtain the following lower bound on t_0.

$$t_0 \;\geq\; \sqrt{\frac{c^2}{4} + \frac{c(L^d - t_0^d)}{dt_0^{d-1}}} + \frac{c}{2},$$

which simplifies to

$$t_0^{d+1} - \left(1 - \frac{1}{d}\right) t_0^d \;\geq\; \frac{cL^d}{d}. \tag{4.2}$$

Since each $p_{d,L}$ is concave, we now invoke inequality (3.14) to obtain the following upper bound on t_0.

$$t_0 \ \le \ 2\sqrt{\frac{c^2}{4} + \frac{c((2L)^d - t_0^d)}{2dt_0^{d-1}}} + c,$$

which simplifies to

$$t_0^{d+1} - 2\left(1 - \frac{1}{d}\right) t_0^d \ \le \ \frac{2^{d+1}cL^d}{d}. \tag{4.3}$$

Using simple estimates based on inequalities (4.2) and (4.3), we find finally that

$$\left(\frac{c}{d}\right)^{1/(d+1)} L^{d/(d+1)} \ \le \ t_0 \ \le \ 2\left(\frac{c}{d}\right)^{1/(d+1)} L^{d/(d+1)} + 1.$$

For the special case $d = 1$, these bounds specialize to

$$\sqrt{cL} \ \le \ t_0 \ \le \ 2\sqrt{cL} + 1, \tag{4.4}$$

which contrasts with the actual optimal value from [3] (stated imprecisely here, to simplify comparison with (4.4)):

$$t_0 \ = \ \sqrt{2cL} + (\text{low-order terms}). \tag{4.5}$$

For specific values of d, we can use ad hoc techniques that emerge from the analysis in Section 5 to get even tighter bounds. Most notably, when $d = 1$, we can revisit (3.1) and the proof of Corollary 5.3, in the light of (4.1) and the fact that $p'_{1,L} \equiv -1/L$, to match (4.5) up to low-order terms.

The ad hoc, but optimal, analysis in [3] builds on the fact that the aggregate overhead from an optimal schedule forms an arithmetic sum.

4.2 The Family $\{p_a(t) \stackrel{\text{def}}{=} a^{-t} \mid a = 1, 2, \ldots\}$

The life functions in this family characterize the *geometrically decreasing lifespan* scenario of [3], which models a cycle-stealing opportunity that has a "half-life."

The non-initial period-lengths. Applying system (3.6) to p_a, we find that the non-initial period-lengths of an optimal schedule $\mathcal{S} = t_0, t_1, \ldots$ satisfy the recurrence[8]

$$a^{-(T_{k-1}+t_k)} \ = \ a^{-T_{k-1}} - (t_{k-1} - c)a^{-T_{k-1}} \ln a,$$

[8]Throughout, "$\ln x$" (resp., "$\log x$") denotes the natural, base-e, (resp., the base-2) logarithm of x.

so that

$$a^{-t_k} + t_{k-1} \ln a \;=\; 1 + c \ln a. \tag{4.6}$$

(Of course, system (4.6) can be solved for $t_1, t_2, \ldots$ only when each $t_k \;<\; c + 1/\ln a$.) System (4.6) contrasts with the actual optimal recurrence

$$a^{-t_k} + t_k \ln a \equiv 1 + c \ln a$$

from [3], wherein the derivation of the recurrence is preceded by a proof that all of the optimal t_k are equal. This proof emerges from the observation that the conditional risk under p_a looks the same at every time instant.

The initial period-length. Lemma 3.1 and inequality (3.7) combine to yield the following bounds on t_0.

$$\sqrt{\frac{c^2}{4} + \frac{c}{\ln a}} + \frac{c}{2} \;\le\; t_0 \;\le\; c + \frac{1}{\ln a}$$

which contrasts with the actual optimal value from [3]:

$$t_0 + \frac{a^{-t_0}}{\ln a} \;=\; c + \frac{1}{\ln a}.$$

Note how close our guidelines' upper bound is to the optimal value.

4.3 The Life Function $p(t) \stackrel{\text{def}}{=} (2^L - 2^t)/(2^L - 1)$

This life function characterizes the *geometrically increasing risk* scenario of [3], which models a cycle-stealing opportunity such as a coffee break, wherein the risk of interruption doubles at every step.

The non-initial period-lengths. Applying system (3.6) to p, we find that the non-initial period-lengths of an optimal schedule $\mathcal{S} = t_0, t_1, \ldots$ satisfy the recurrence

$$2^L - 2^{T_{k-1}+t_k} \;=\; 2^L - 2^{T_{k-1}} - (t_{k-1} - c)2^{T_{k-1}} \ln 2,$$

so that

$$t_{k+1}^{(L)} \;=\; \log((t_k^{(L)} - c) \ln 2 + 1). \tag{4.7}$$

System (4.7) contrasts with the actual optimal recurrence from [3]:

$$t_{k+1}^{(L)} \;=\; \log(t_k^{(L)} - c + 2),$$

which emerges from comparing each $\mathcal{S}$ with its *kth-period perturbations* (cf. Section 5.1):

$$\mathcal{S}^{\pm k} \stackrel{\text{def}}{=} t_0, t_1, \ldots, t_{k-1}, t_k \pm 1, t_{k+1} \mp 1, t_{k+2}, \ldots, t_{m-1}.$$

The initial period-length. Since the function p is concave, we invoke inequalities (3.7) and (3.14) to yield, respectively, lower and upper bounds on t_0. Without writing out the long expressions that these inequalities yield, we note only that they show that, to within low-order additive terms (which involve c, t_0, and L),

$$2^{t_0/2} t_0^2 \;\le\; 2^L \;\le\; 2^{t_0} t_0^2.$$

It follows that, to within low-order additive terms,

$$t_0 \;=\; \frac{L}{\log^2 L}.$$

No explicit value for t_0 is derived in [3].

5 Further Insights into Optimal Schedules

This section is devoted to establishing a number of properties of optimal schedules, which, particularly for concave life functions: lend insight into the nature of optimal schedules; sometimes help one evaluate the "formulas" (3.1) for optimal period-lengths; sometimes help one sharpen the bounds on t_0 from Theorems 3.2 and 3.3.

5.1 The "Local" Sufficiency of the Inter-Period Dependencies

Theorem 3.1 fails to *characterize* optimal schedules because its system of dependencies is shown only to be necessary for optimality. We do not presently have a proof of the sufficiency of system (3.1), but we can take a small step in the direction of such a proof. Specifically, we prove now that system (3.6) guarantees the "local" optimality of a schedule, at least for concave life functions.

Our informal notion of the "local" optimality of schedule $\mathcal{S}$ builds on the formal notion of a "perturbation" of $\mathcal{S}$ (which supplements the earlier notion of a "shift" of $\mathcal{S}$).

> The $[k, -\delta]$-*perturbation*, $\mathcal{S}^{[k,-\delta]}$, of $\mathcal{S}$ and the $[k, +\delta]$-*perturbation*, $\mathcal{S}^{[k,+\delta]}$, of $\mathcal{S}$ are the schedules
>
> $$\begin{array}{lcl} \mathcal{S}^{[k,-\delta]} & \stackrel{\text{def}}{=} & t_0, t_1, \ldots, t_{k-1}, t_k - \delta, t_{k+1} + \delta, t_{k+2}, \ldots \\ \mathcal{S}^{[k,+\delta]} & \stackrel{\text{def}}{=} & t_0, t_1, \ldots, t_{k-1}, t_k + \delta, t_{k+1} - \delta, t_{k+2}, \ldots \end{array}$$

which have the same number of periods as $\mathcal{S}$ and the same period-lengths, save for periods k and $k+1$.

Theorem 5.1 *Let $\mathcal{S} = t_0, t_1, \ldots$ be a schedule for a* concave *life function p. If the period-lengths of $\mathcal{S}$ satisfy system (3.6), then schedule $\mathcal{S}$ is more productive than any of its δ-perturbations; i.e., for all period-indices i and real $\delta > 0$, $E(\mathcal{S}; p) > E(\mathcal{S}^{[i,-\delta]}; p)$, and $E(\mathcal{S}; p) > E(\mathcal{S}^{[i,+\delta]}; p)$.*

Proof. **(a)** Note first that, if $\mathcal{S}$ satisfies system (3.6), then, for each period-index i,

$$p(T_{i+1}) \;<\; p(T_i - \delta) + (t_i - c)\left[\frac{p(T_i) - p(T_i - \delta)}{\delta}\right] \tag{5.1}$$

for all $\delta > 0$. This follows from the $(j = i)$ instance of system (3.6), coupled with the following two facts. First, $p(T_i) < p(T_i - \delta)$ because p is a decreasing function. Second, by the Mean-Value Theorem of the Differential Calculus (henceforth, "the MVT" for short), the fraction in (5.1) equals $p'(\xi)$ for some $\xi \in (T_i - \delta,\ T_i)$. Since $\xi < T_i$ and since p is concave, we have $p'(T_i) \le p'(\xi)$.

Now, for each i and δ, we can manipulate the relevant instance of system (5.1) to obtain:

$$[(t_i - \delta - c)p(T_i - \delta) + (t_{i+1} + \delta - c)p(T_{i+1})] - [(t_i - c)p(T_i) + (t_{i+1} - c)p(T_{i+1})] < 0. \tag{5.2}$$

Since the lefthand side of (5.2) equals $E(\mathcal{S}^{[i,-\delta]}; p) - E(\mathcal{S}; p)$, and since i and δ were arbitrary, we conclude that $E(\mathcal{S}; p) > E(\mathcal{S}^{[i,-\delta]}; p)$ for all i and δ.

(b) We now mimic the reasoning in part (a), using positive perturbations and concave life functions. To wit, if schedule $\mathcal{S}$ satisfies system (3.6), then, for each period-index i,

$$p(T_{i+1}) \;>\; p(T_i + \delta) + (t_i - c)\left[\frac{p(T_i + \delta) - p(T_i)}{\delta}\right] \tag{5.3}$$

for all $\delta > 0$. In parallel with part (a), this follows from the $(j = i)$ instance of system (3.6) in the presence of: the fact that p is a decreasing function; an invocation of the MVT, coupled with an invocation of p's concavity. Now, as before, we manipulate the relevant instance of system (5.3) to obtain

$$[(t_i - c)p(T_i) + (t_{i+1} - c)p(T_{i+1})] - [(t_i + \delta - c)p(T_i + \delta) + (t_{i+1} - \delta - c)p(T_{i+1})] > 0. \tag{5.4}$$

Since the lefthand side of (5.4) is just $E(\mathcal{S}; p) - E(\mathcal{S}^{[i,+\delta]}; p)$, and since both i and δ were arbitrary, we conclude that $E(\mathcal{S}; p) > E(\mathcal{S}^{[i,+\delta]}; p)$ for all i and δ. ■

5.2 The Growth Rate of Optimal Period-Lengths

We now present a result that establishes rather weak bounds on the growth rate of the period-lengths of optimal schedules. Specifically, when the life function p is concave (resp., convex), then each "internal" period—i.e., excepting the last one—should be *at least* (resp., *at most*) c time units longer than its successor. Despite their weakness, these bounds have analytically useful consequences, most particularly proving that optimal schedules for concave life functions must be finite and providing upper bounds on their numbers of periods.

Theorem 5.2 *Let $\mathcal{S} = t_0, t_1, \ldots$ be an optimal schedule for a cycle-stealing episode with life function p. If p is* concave*, then each $t_{i+1} \leq t_i - c$. If p is* convex*, then each $t_{i+1} \geq t_i - c$.*

Proof. We actually exploit the optimality of $\mathcal{S}$ only to infer that it is at least as productive as any of its δ-perturbations. Our arguments for concave and convex p are very similar, but they differ in essential technicalities.

(a) Consider first the case of concave p. Since schedule $\mathcal{S}$ is optimal for p, we know that $E(\mathcal{S}; p) \geq E(\mathcal{S}^{[i,+\delta]}; p)$ for every period-index i and every real $\delta > 0$. Focusing on fixed but arbitrary i and δ, we find that

$$E(\mathcal{S}; p) - E(\mathcal{S}^{[i,+\delta]}; p) \;=\; (t_i - c)\,[p(T_i) - p(T_i + \delta)] + \delta\,[p(T_{i+1}) - p(T_i + \delta)] \;\geq\; 0.$$

We infer that

$$\left(\frac{t_{i+1} - \delta}{t_i - c}\right) \frac{p(T_{i+1}) - p(T_i + \delta)}{t_{i+1} - \delta} \;\geq\; \frac{p(T_i + \delta) - p(T_i)}{\delta}. \tag{5.5}$$

By the MVT, there exist real numbers $\xi \in (T_i, T_i + \delta)$ and $\eta \in (T_i + \delta, T_{i+1})$ such that

$$p'(\xi) = \frac{p(T_i + \delta) - p(T_i)}{\delta} \quad \text{and} \quad p'(\eta) = \frac{p(T_{i+1}) - p(T_i + \delta)}{t_{i+1} - \delta}. \tag{5.6}$$

Now, the concavity of p implies that $p'(\xi) \geq p'(\eta)$, because $\xi < \eta$. Since p' is negative, this inequality can coexist with (5.5) and (5.6) only if $t_{i+1} - \delta < t_i - c$. Since this last inequality holds for each i and for arbitrarily small δ, we conclude that each $t_{i+1} \leq t_i - c$.

(b) Consider next the case of convex p. Since schedule $\mathcal{S}$ is optimal for p, we know that $E(\mathcal{S}; p) \geq E(\mathcal{S}^{[i,-\delta]}; p)$ for every period-index i and every real $\delta > 0$. Focusing on fixed but arbitrary i and δ, we find that

$$E(\mathcal{S}; p) - E(\mathcal{S}^{[i,-\delta]}; p) \;=\; (t_i - c)\,[p(T_i) - p(T_i - \delta)] - \delta\,[p(T_{i+1}) - p(T_i - \delta)] \;\geq\; 0.$$

We now infer (after some manipulation) that

$$\frac{p(T_i) - p(T_i - \delta)}{\delta} \;\geq\; \left(\frac{t_{i+1}}{t_i - \delta - c}\right) \frac{p(T_{i+1}) - p(T_i)}{t_{i+1}}. \tag{5.7}$$

The remainder of the proof mimics part (a). After two invocations of the MVT and one of the convexity of p, we end up with the inequality

$$p'(\xi) \geq \left(\frac{t_{i+1}}{t_i - \delta - c}\right) p'(\eta),$$

where $p'(\xi) \leq p'(\eta) < 0$. Clearly, this inequality on p' can coexist with (5.7) only if $t_{i+1} + \delta > t_i - c$. We now conclude, as in part (a), that each $t_{i+1} \geq t_i - c$. ■

The analysis in [3] of uniform-risk life functions shows that Theorem 5.2 cannot be improved in general: each such function is both concave and convex, and the period-lengths of its unique optimal schedule satisfy $t_{i+1} = t_i - c$ for all i (see Section 4.1).

Theorem 5.2 tells us quite a bit about optimal schedules for concave life functions. Firstly, it tells us that we should select period-lengths that are strictly decreasing. This strengthens an analogous result in [3], which is proved there only with weak inequalities and only for the uniform-risk scenario.

Corollary 5.1 *If the schedule $\mathcal{S} = t_0, t_1, \ldots$ is optimal for a* concave *life function, then, for each period-index i, $t_i > t_{i+1}$.*

Next, Theorem 5.2 tells us that optimal schedules for concave life functions are finite, and it yields bounds on their numbers of periods.

Corollary 5.2 *An optimal schedule $\mathcal{S} = t_0, t_1, \ldots$ for a* concave *life function is finite, having no more than t_0/c periods.*

Proof. By Theorem 5.2, $\mathcal{S}$'s period-lengths decrease at the rate of at least c per period; by Proposition 2.1, all of $\mathcal{S}$'s periods, save the last, have length $> c$. ■

Neither Corollary 5.1 nor 5.2 is true in general. Specifically, the unique optimal schedule for the geometrically decreasing lifespan scenario is infinite and has all period-lengths equal [3] (see Section 4.2).

Since optimal schedules for concave life functions are finite, and since each period of any schedule has finite duration, it follows that every cycle-stealing episode that is characterized by a concave life function has a bounded potential lifespan L. The knowledge that such an L exists enables us to fine-tune Corollary 5.2.

Corollary 5.3 *Let $\mathcal{S}$ be an optimal schedule for a* concave *life function with potential lifespan L. The number m of periods of $\mathcal{S}$ satisfies*

$$m \ < \ \left\lceil \sqrt{\frac{2L}{c} + \frac{1}{4}} + \frac{1}{2} \right\rceil . \tag{5.8}$$

Proof. When one looks at Theorem 5.2 "from the vantage point of t_{m-1}," one finds that

$$\begin{aligned} L &= t_0 + t_1 + \cdots + t_{m-2} + t_{m-1} \\ &\geq m t_{m-1} + \binom{m}{2} c \\ &> \binom{m}{2} c. \qquad \blacksquare \end{aligned} \tag{5.9}$$

The analysis in [3] of the uniform-risk scenario shows that the bound of Corollary 5.3 cannot be improved in general. Specifically, number of periods of the unique optimal schedule for $p_L = 1 - t/L$ is given by (5.8) with floors replacing ceilings.

Finally, Theorem 5.2 supplements Theorems 3.3 and 3.2 with additional bounds on the optimal value of t_0.

Corollary 5.4 *Let $\mathcal{S} = t_0, t_1, \ldots$ be an optimal schedule for life function p. If p is* concave*, and the cycle-stealing episode has potential lifespan L, and schedule $\mathcal{S}$ has m periods, then*

$$t_0 \ \geq \ \frac{L}{m} + \frac{m-1}{2} c. \tag{5.10}$$

Proof. The bound (5.10) follows from looking at Theorem 5.2 "from the vantage point of t_0":

$$\begin{aligned} L &= t_0 + t_1 + \cdots + t_{m-2} + t_{m-1} \\ &\leq m t_0 - \binom{m}{2} c \qquad \blacksquare. \end{aligned}$$

Corollary 5.5 *If the life function p is* concave *and has potential lifespan L, then the optimal t_0 satisfies:*

$$t_0 \ > \ \sqrt{\frac{cL}{2}} + \frac{3}{4} c \quad \textit{and} \quad p(t_0) \ > \ -\frac{1}{8} \left(L - \sqrt{\frac{c}{8}} \right)^2 p'(t_0).$$

Proof. The lefthand inequality is immediate from (5.10) in the presence of Corollary 5.3. The implicit bound of the righthand inequality follows by composing (3.13) with the lefthand inequality. $\blacksquare$

6 Conclusions

Our experience with specific life functions, as illustrated in Section 4, suggests that, despite its implicit nature, system (3.6) easily determines each non-initial period-length of an optimal schedule in terms of all earlier period-lengths. Significantly, this "progressive" feature of the system allows one to determine t_{i+1} only after period i has ended. This means that, in principle, one could use *conditional*, rather than absolute, probabilities to determine schedule $\mathcal{S}$ progressively, period by period. Determining the initial period-length t_0 remains an art. System (3.6) does not help in this determination, and the $(k = 0)$-instance of system (3.1) is usually hard to apply, except in very special cases, such as the uniform risk scenario. The bounds on the optimal value of t_0 that we derive in Theorems 3.2 and 3.3 substantially narrow one's search space for the optimal t_0, at least for "smooth" life functions, but they usually still leave one with a factor-of-2 uncertainly in determining this value. Indeed, we view the primary open problem within the framework we have studied here to be the identification of broad classes of life functions for which one can determine the optimal initial period-length t_0. Although we succeeded in [3] to make this determination for three specific scenarios, the techniques used there were very specific to the particular life functions being studied.

Even aside from determining t_0 definitively, many basic questions remain within the framework of our study. Most obviously, our results expose only *necessary* dependencies among optimal period-lengths; they do not demonstrate that using such period-lengths guarantees the (near) optimality of the resulting schedule. (Of course, Theorem 5.1 is a step toward filling this gap.) While avenues toward global optimality guarantees have eluded us, one possible approach would involve answering the following question.

Are optimal cycle-stealing schedules unique?

Significantly, Theorem 3.1 gives a handle on this basic question, since it implies that distinct optimal schedules must have different *initial* period-lengths. Notably, each of the life functions studied in [3] admits a unique optimal schedule—but the techniques for verifying uniqueness there were specific to the individual life function. Yet another approach to guaranteeing optimality—at least for specific classes of life functions—would be to determine when specific scheduling recipes work. One natural such recipe is to choose period-lengths "greedily:" one would choose t_0 by maximizing the function $p_0(t) \stackrel{\text{def}}{=} (t-c)p(t)$, then choose t_1 by maximizing the function $p_1(t) \stackrel{\text{def}}{=} (t-c)p(t+t_0)$, and so on.

For what class of life functions is a "greedy" cycle-stealing schedule optimal?
In general, how good are "greedy" schedules?

Easily, the "greedy" strategy yields the optimal schedule for the geometrically decreasing lifespan scenario—but it does not for the uniform-risk scenario. In quite another direction, we do not yet have an answer to even the following basic question, whose nontriviality is attested to by Corollary 3.2.

For what class of life functions do there exist optimal cycle-stealing schedules?

A final set of open questions involve more technical issues. Our current results demand smoothness and/or a nice "shape" in our life functions. Can these assumptions be weakened? In another direction: we have had to translate what is ideally a discrete problem into a continuous framework in order to derive our guidelines; this was true even in the case study of [3]. Can one show that our continuous guidelines yield valuable discrete analogues?

It is clear from this brief list of questions that many challenges remain in this important area of research.

Acknowledgments. This research was supported in part by NSF Grant CCR-97-10367. It is a pleasure to thank Ramesh Sitaraman for being a patient and helpful audience as I was doing this research.

References

[1] M.J. Atallah, C.L. Black, D.C. Marinescu, H.J. Siegel, T.L. Casavant (1992): Models and algorithms for coscheduling compute-intensive tasks on a network of workstations. *J. Parallel Distr. Comput. 16*, 319-327.

[2] B. Awerbuch, Y. Azar, A. Fiat, F.T. Leighton (1996): Making commitments in the face of uncertainty: how to pick a winner almost every time. *28th ACM Symp. on Theory of Computing*, 519-530.

[3] S.N. Bhatt, F.R.K. Chung, F.T. Leighton, A.L. Rosenberg (1997): On optimal strategies for cycle-stealing in networks of workstations. *IEEE Trans. Comp. 46*, 545-557.

[4] R. Blumofe and C.E. Leiserson (1993): Space-efficient scheduling of multithreaded computations. *25th ACM Symp. on Theory of Computing*, 362-371.

[5] R. Blumofe and C.E. Leiserson (1994): Scheduling multithreaded computations by work stealing. *35th IEEE Symp. on Foundations of Computer Science*, 356-368.

[6] R. Blumofe and D.S. Park (1994): Scheduling large-scale parallel computations on networks of workstations. *3rd Intl. Symp. on High-Performance Distributed Computing*, 96-105.

[7] E.G. Coffman, Jr., L. Flatto, A.Y. Krenin (1993): Scheduling saves in fault-tolerant computations. *Acta. Inform. 30*, 409-423.

[8] T. Hagerup (1998): Allocating independent tasks to parallel processors: an experimental study. *J. Parallel Distr. Comput.*, to appear.

[9] C.H. Papadimitriou and M. Yannakakis (1990): Towards an architecture-independent analysis of parallel algorithms. *SIAM J. Comput. 19*, 322-328.