A Decision Process Analysis of Recently Reported Implicit CO-scheduling Results

R. Poovendran

ABSTRACT

Using a series of elaborate simulation,[?,?,?] Andrea et al proposed a coscheduling algorithm based on two-phase spin blocking. By performing benchmark measurements, they derived a set of systems parameters and using heuristic arguments, they showed how a local processor can make the decision to either spin the process for some more time or to block. One of their result was based on the barrier imbalance and the other one was based on pairwise cost-benefit. They also noted that their cost derivations were for competing but identical processes that needed to be coscheduled. We note that it is possible to model the experimental setup as a binary Bayesian decision process at the local work station and choosing appropriate parameters lead to the results derived in.[?,?,?] Hence, the results derived are interpretable using well founded theoretical framework that not only only interpreted the results, may also help in studying the variations they could have obtained by analysis before experiments.

1. Introduction

1.1. Background on Binary Decisions Problem

We first describe a binary detection problem and map it to the DSM problem. Then we present the optimal decision making based on Bayes criteria.

In a binary decision problem, the observer (local processor) makes a set of decisions based on the possible two final outcomes of the problem. One of the outcomes contains some important relevant data or message and the other one contains no important data. The outcome that contains no relevant data is called the *null hypothesis* (H_0) and the other one is denoted as *alternative hypothesis* (H_1). Each hypothesis contains one or more observations which are represented by random variables. When the observer makes the decision, depending on the final actual outcomes, the following four cases can occur.

1. H_0 was decided and H_0 occurred.
2. H_0 was decided and H_1 occurred.
3. H_1 was decided and H_0 occurred.
4. H_1 was decided and H_1 occurred.

We note that the observer makes correct decisions in cases (i) and (iv), while the observer makes wrong decisions in cases (ii) and (iii). In terms of the process scheduling we can decide the set H_1 as the event that there is a coordination or message or barrier synchronization from a remote host and H_0 as no such possibility. Depending on the binary decisions, it can be inferred whether it will be cost effective to let the local process spin or go to sleep.

The four decision outcomes can be related to inferences as:

1. Decide it is better to block the process and save unnecessary spinning.
2. Decide it is better to block the process and pay a penalty by missing the message.
3. Decide it is better to spin and pay a penalty by wasting resources.
4. Decide it is better to spin and benefit by being able to respond to the incoming message.

We note that the case (ii) is related to a *miss* and the case (iv) is related to a *false alarm* in the context of decision theory.

1.2. Bayesian Criterion

In using Bayesian criteria, two assumptions are made. First, the probability of the occurrences of two outcomes are known. They are the probabilities $P_0 = P(H_0)$ and $P_1 = P(H_1)$. We note that

$$P_0 + P_1 = 1 \tag{1}$$

Each decision has a cost associated with it. For example, if the local process was allowed to spin and no message was received or barrier was completed, then there is a wasted time that could have been assigned to some other process. Similarly, if the local process was blocked and had to be woken up then there is an associated cost for missing the synchronization or coscheduling. Clearly, for a simple binary decision problem, there is no cost associated with making the right decisions. However, in a composite hypothesis problem it may matter (as in the case of fairness across multiple type of process reported i.[?])

Let $D_i; i = 0, 1$ where D_0 denotes "decide H_0" and D_1 denotes "decide H_1", we can define C_{ij} as the cost associated with the decision D_i given that the true hypothesis is H_j. In particular, the cost assignment for all the four cases of the binary hypothesis are given below

1. C_{00} for case 1
2. C_{01} for case 2
3. C_{10} for case 3
4. C_{11} for case 4

We also assume that the cost of making a wrong decision is more than cost of making a correct decision. That is,

$$C_{01} > C_{11} \tag{2}$$
$$C_{10} > C_{00} \tag{3}$$
$$\tag{4}$$

Given $(P(D_i, H_j)$, the joint probability that we decide D_i and the hypothesis H_j is true, the average cost/risk is given by

$$R = E[C] = C_{00}P(D_0, H_0) + C_{01}P(D_0, H_1) + C_{10}P(D_1, H_0) + C_{11}P(D_1, H_1) \tag{5}$$

From Bayes rule we have $P(D_i, H_j) = P(D_i|H_j)P(H_j)$. If we denote that the underlying distributions of hypotheses H_0 and H_1 are given by f_{h0} and f_{h1}, after some algebra the decision rule comes to

$$\frac{f_{h0}}{f_{h1}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \tag{6}$$

Since the prior distributions are known, if we denote

$$\lambda(y) = \frac{f_{h0}}{f_{h1}} \tag{7}$$
$$\eta = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \tag{8}$$

Then, the Bayes criteria reduces to

$$\lambda(y) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{9}$$

We will now intepret the results of the experiments reported in.[?] We denote the cost of blocking as a function of W as $f(W)$ instead of a constant.

1.3. Interpretation of the results in?: Local Cost-Benefit

One of the main decision made in? was the criterion for deciding whether to spin for more time in the presence of load-imbalance or not. In that experiment, the authors ran similar processes that were competing to be coscheduled. Since the processes are identical, their underlying distributions were identical. Hence, at any given time, there was equal probability of receiving message from any one of the competing process set. That is, the message, or barrier synchronization did not have any bias towards any particular process. Hence, $f_{h0} = f_{h1}$ was assumed in the experiments. (We note that this is not really a case in heterogeneous real world applications) Moreover, the simulations in? assign equal probalities to receiving or not receiving a message or load-imbalance. This reduces to $P_0 = P_1 = 0$ in the Bayesian decision.

From their experiments, the load-imbalance cost functions are (note that the penalty is a function proportional to time spinning in their study as it should be)

1. $C_{00} = 0$ (i.e. if blocking was correct no penalty is payed)
2. $C_{01} = (B + f(W)) + f(W)$. i.e. penalty for blocking is the spin time $B + f(W)$ and then the penalty $f(W)$ for waking up.
3. $C_{10} = f(V)$. i.e. the penalty for spinning to wait for the barrier synchronization without reaching it.
4. $C_{11} = 0$. i.e. pay no penalty by being spinning since the process arrives on time.
5. In the simulations reported in,? $f(W) = W$ and $f(V) = V/2$ were assumed. We note that these two quantities probably can be better measured using a benchmark. with these quantities, the decision problem reduces to

$$1 \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \frac{(C_{10} - 0)}{(C_{01} - 0)} \tag{10}$$

$$\Rightarrow C_{01} \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad C_{10} \tag{11}$$

In this case, their decision process reduces to

$$Decide \begin{cases} \text{spin wait} & \text{if } (C_{10} = V/2) < (C_{01} = B + 2W) \\ \text{block} & else \end{cases}$$

In simple words, if $v < 2(B + 2W)$ it is beneficial to spin waiting for the load-imbalance to close.

1.4. Pairwise: Cost-Benefit: Incoming Messages

In the second experiment, authors developed their results based on the intuition that a process handling messages should continue to spin. H_0 here represent the event that there is no-messageandH$_1$ here is message arrival. As in the previous case, identical but competing jobs were used for coscheduling. Probabilities on receiving or not receiving a message were set equal implicitly. The correct decisions were not assigned any penalties. Round trip and the processing overhead at the end-to-end pairwise processing time was $2L + 4O$. The costs, derived from various time requirements are enumerated below:

(a) $C_{00} = 0$ (i.e. if blocking was correct no penalty is payed)

(b) $C_{01} = 2L + 4O + 5W$. i.e. penalty for blocking is the flight latency, processing time, "short spin" time and three more block related penalties.

(c) $C_{10} = 2L + 4O + T$. i.e. the penality for latency + processing + additional wait.

(d) $C_{11} = 0$. i.e. pay no penalty by being spinning since the process arrives on time.

With these quantities, the Bayes decision reduces to the their decision process as

$$Decide \begin{cases} \text{spin wait} & \text{if } (C_{10} = 2L + 4O + T) < (C_{01} = 2L + 4O + 5W) \\ \text{block} & else \end{cases}$$

This leads to the result obtained by the authors? as decide to spin more time if $T < 5W$.

2. Discussions

We note that the scope of the experiment is limited in the following sense.

(a) Authors assumed that the competing jobs are all identical.

(b) They also assumed that the probability of receiving or not receiving a request for a particular process is identical.

(c) They assumed that there is no penalty for making the right decision with respect to one process.

(d) They assumed that W is already known fixed quantity.

However, from their later experiments, it becomes clear that they found that when the competing processes are not identical, they could not provide any derivations. If they had realize the decision model they were using, they could have assigned non-identical values to different processes and demonstrated the analytical values against the measured values. This, in my opinion plays a role in the heterogeneous environment.