Referee Report

Paper Number:
Reviewer's Name: Jesse Bannon
Name of Paper: X: A Comprehensive Analytic Model for Parallel Machines
Author(s): Ang Li, Shuaiwen Leon Song, Akash Kumar, Daniel Chavarria-Miranda, Henk Corporaal

Section I. Overview
A. Reader Interest

1. Which category describes this manuscript?

( ) Practice / Application / Case Study / Experience Report
(x) Research / Technology
( ) Survey / Tutorial / How-To

2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under IIIA.

( ) Very Relevant
(x) Relevant
( ) Interesting - but not very relevant
( ) Irrelevant

B. Content

1. Please explain how this manuscript advances this field of research and / or contributes something new to the literature. Please explain your answer under IIIA. Public Comments.

2. Is the manuscript technically sound? Please explain your answer under IIIA. Public Comments.

( ) Yes
(x) Appears to be - but didn't check completely
( ) Partially
( ) No

C. Presentation

1. Are the title, abstract, and keywords appropriate? Please explain your answer under IIIA. Public Comments.

(x) Yes
( ) No

2. Does the manuscript contain sufficient and appropriate references? Please explain your answer under IIIA.

(x) References are sufficient and appropriate
( ) Important references are missing; more references are needed
( ) Number of references are excessive

3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under IIIA. Public Comments.

( ) Yes

(x) Could be improved
( ) No

4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain your answer under IIIA. Public Comments.

(x) Satisfactory
( ) Could be improved
( ) Poor

5. Please rate and comment on the readability of this manuscript. Please explain your answer under IIIA.

(x) Easy to read
( ) Readable - but requires some effort to understand
( ) Difficult to read and understand
( ) Unreadable

Section II. Summary and Recommendation

A. Evaluation

Please rate the manuscript. Please explain your answer under IIIA. Public Comments.

( ) Award Quality
( ) Excellent
(x) Good
( ) Fair
( ) Poor

B. Recommendation

Please make your recommendation. Please explain your answer under IIIA. Public Comments.

( ) Accept with no changes
(x) Author should prepare a minor revision
( ) Author should prepare a major revision for a second review
( ) Reject

Section III. Detailed Comments

A. Public Comments (these will be made available to the author)

Explanation for the Recommendation

The paper is well organized and rich in content. There are three items missing from making this a complete paper.

The first is the describing the importance of cache in the X-Model. The abstract does not mention cache, and the introduction barely touches on it. Yet there are two subsections in Section III devoted to cache effects. Cache should be just as relevant as the types of parallelisms in both the abstract and introduction.

What lacks is reproducibility instructions to replicate the findings shown in Figure 11. In Section V, it is unclear how some of the parameters, notably E, are derived. More information could be presented on how exactly assembly

code is parsed to determine this value. The parameter Z is more well defined, but for the sake of results, it would be helpful to display these values for each of the applications.

Another value that is unclear is how the prediction accuracy was obtained in Section V. The graphs show the star being close to the intersection of nearly every graph, but this exact value's derivation is unclear.

Summary of the Paper and Assessment

This paper present a new metric for single-node parallel systems. While it doesn't say explicitly, the metric is for data-parallel tasks. The introduction and abstract state what is missing from standard procedures when measuring these kinds of problem's performance, and how X-Model takes into account these missing considerations.

Section II immediately informs the reader that the work presented in this paper is a continuation their previous Transit Model, and describes the differences between the two in detail. This is a great introductory section describing some of the elementary factors before diving into the innards of the X-Model. The table defining all of the parameters was especially useful for later sections when describing equations and properties of the plots.

Section III ties back into the introduction of missing properties when measuring performance, by describing in depth how the new parameters introduced address these problems, and how they directly relate to the overall performance of a program. These properties include Memory-Level-Parallelism (MLP), Instruction-Level-Parallelism (ILP), Thread-Level-Parallelism (TLP), Data-Level-Parallelism (DLP), and the effects of cache. The parameter table in Section II makes this section very easy to read. Section III devotes two entire subsections to cache effects. The relevance of cache in this paper is not made clear in the abstract or the introduction, yet is described as one of the key driving factors in the X-Model.

Sections IV and VI reiterate the selling point of how the X-Model requires very few parameters when conducting performance analysis, and how to select these parameters using their validation results as an example. Some parameters, notably E and Z, and vague in their descriptions on how to obtain them exactly.

Section V shows their metric and practices described in Section IV on real-world programs, and how closely their model relates to run-time metrics taken by the CUDA profiler. Their claim is 84.1% prediction accuracy, but how that number was obtained is not clear.

Section VI compares the X-Model to three different performance models. Each model measures a subset of what the X-Model measures, which can be viewed as either an unfair comparison or the X-Model is so rich in features, that nothing else compares to it. I do not know enough on the subject to determine this. Regardless, the comparison aides the X-Model's features when measuring parallel systems.