

# Proposal Summary

## 1. Identification

Program Solicitation: #NSF 04-012

Proposal title: ITR-(ASE)-(int): Advancing machine performance by data-driven optimization models with back-tracking mechanism

Name of PI: Xiaoya Xiang

Department: URCS

## 2. Main idea of the Research

Through the research, we mean to discover the general rules and patterns of information interaction between optimizations, so as to find out a nearly-best optimization sets on existing machine and compiler.

1) Set up a model driven by data affinity to optimize program and machine performance. This model takes into account various aspects of machine's performance such as cache behavior, instruction scheduling, register allocation and loop overhead, and estimates the overall performance.

2) Adding effective back-tracking mechanism to the model to exchange information in run time and to revise the strategy of the model to maximize performance gains.

## 3. Relationship To Other Research & Practice

The idea of data-driven optimizations is not new. Ding used a sequential scheme to explore a two-step strategy [1, 2], which is most similar to the first part of our research. However, his work did not consider the effect on the following-after optimizations and the back-tracking mechanism.

To date, the most popular method for information interchanging is iterative method. Many people focus on this topic and have produced many works (e.g. [3]). However, it

seems iterative method is not quite suitable to our model since it assumes the data layout be not changeable. As a result, our mechanism may be essentially different from the iterative ones.

reference:

[1] C. Ding and K. Kennedy. Inter-array data regrouping. In Proceedings of The 12th International Workshop on Languages and Compilers for Parallel Computing, La Jolla, California, August 1999.

[2] C. Ding and K. Kennedy. Improving effective bandwidth through compiler enhancement of global cache reuse. Journal of Parallel and Distributed Computing, 64(1), 2004.

[3] Grigori Fursin and Albert Cohen. Building a Practical Iterative Interactive Compiler. To appear at the 1<sup>st</sup> Workshop on Statistical and Machine Learning Approaches Applied to Architectures and Compilation (SMART'07), Ghent, Belgium, January 2007

## 4. Research Potential Impact

The passed decades have seen a stream of research on understanding the pattern of data access and to what degree that pattern can be modeled, measured and improved. The strategies, which make some changes to the data layout to make full use of data locality and affinity, works especially well when evaluated by cache behavior. But nothing comes free. Aggressive optimizations often inhibit later transformations, and consequently reduce performance gains or even hurt performance,

Our research mainly targets the information interaction between optimizations. Using data driven optimizations and sufficient back-tracking information, we aim at make best use of data locality and affinity while maintaining or even improving other optimizations.

Main potential benefits involves:

- 1) Improving program performance on the given machine.
- 2) Alleviate Memory Wall and Power Wall, both of which is the main problems faced by modern computer architecture, especially multi-core or many-core architecture. Since data is well arranged and used, CPU performance bottleneck is alleviated. Even power is saved.

Additional benefits contains but not restrict to: providing an alternative information interacting mechanism other than iteration, and a new evaluating model to check up the contributions of various optimizations; providing suggestions to hardware design and

programming model design.

## 5. Research Plan

The outline of research plan is as follow:

- deliverables (e.g., demos, videos, prototypes, papers, seminars)
  - The implementation of the data-driven model in the existing compiler
  - An effective back-tracking mechanism added to the above model
  - A new model for evaluation of optimizations
  - A complete technical report on the implementation and performance analysis
  - 1-2 papers will be published on major international conferences.

- personnel

-Xiaoya Xiang

-AAA, BBB, CCC (any person who are interested in this proposal... )

- milestones (it is only assumed to be like this ...)

-April, 2009, Implementation of data-driven model.

-June, 2009, the mechanism of back-tracking.

-November 2007, Evaluation, comparison and analysis

- technical tradeoffs that may have to be made

The data-driven optimizations always change the data layout and code layout, which inhibit later transformations and consequently reduce performance gains or even hurt performance, Furthermore, the back tracking is usually time-consuming and involves big mass of information and complexity relationships between these information. We should carefully handle these issues to prevent from heavy overhead and worse performance.

- proposed plan for technology exchange and interaction with the public

A quarterly update on progress and plans will be reported to the webpage. Any work published as a result of this research will be notified too. The code will be open-sourced when it is finished.

## 6. Budget Detail

N/A

## 7. Resources

- The open-sourced ORCC compiler
- Valgrind machine instrumenting tools (available from Internet for free!!)
- Multi-core computers with details of the architecture.
- performance simulation and validation tools. such as sim-cache ....

## 8. Personal Data

[www.cs.rochester.edu/~xiang](http://www.cs.rochester.edu/~xiang)