Towards Automatic Cache Performance Analysis and Optimization

Tanzima Islam[1], Allyson Cauble-Chantrenne[2], Kathryn Mohror[3], Barry Rountree[3]
[1]Purdue University, [2]U. C. San Diego, [3]Lawrence Livermore National Lab

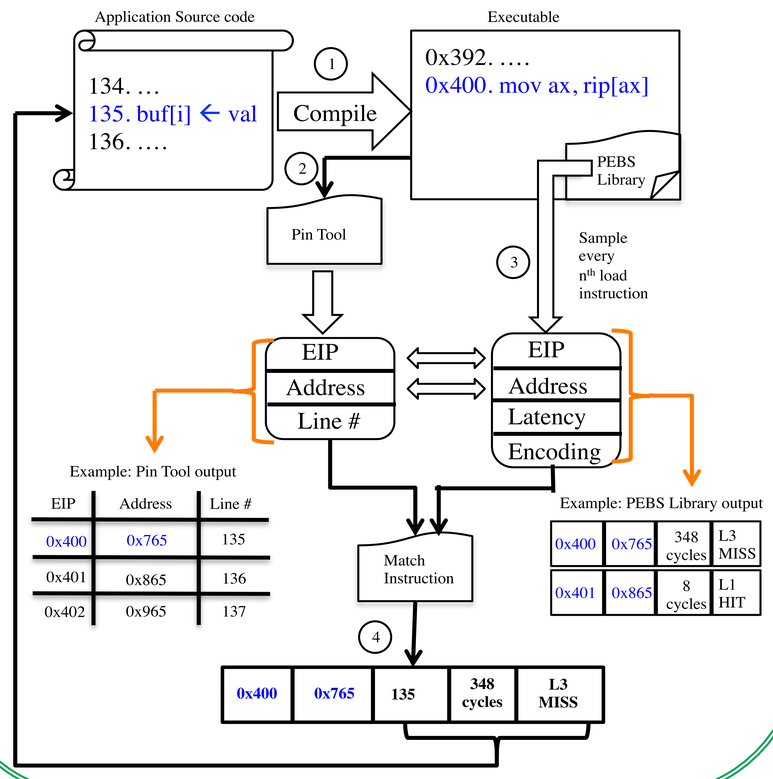
The complexity of today's computing systems greatly increases the difficulties of tuning applications for optimal memory performance. Because currently available tools do not give information about the correlation between cache usage and application performance, developers resort to time-consuming, non-portable "hand-tunings" for complex scientific applications. To address this challenge, we present a user-space tool that makes the first step towards automating the process of optimizing application cache layout and utilization.

Goal

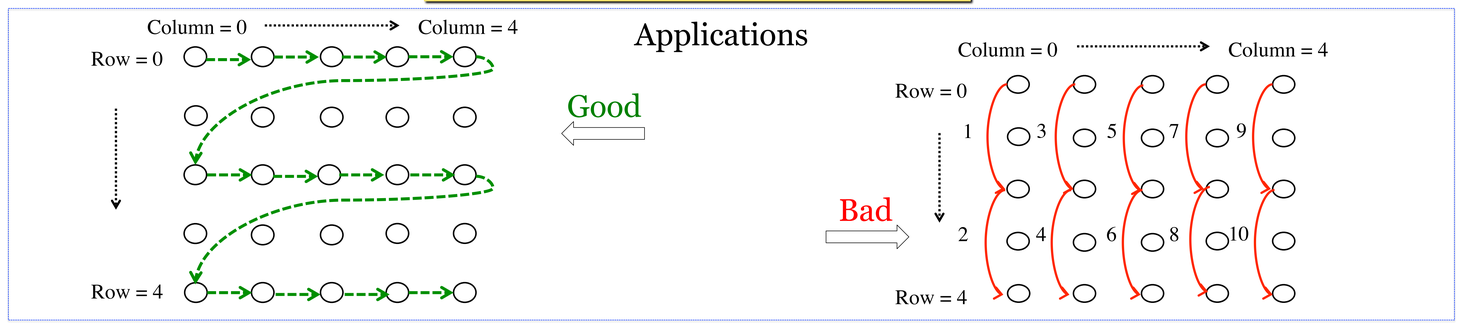
To study application cache performance

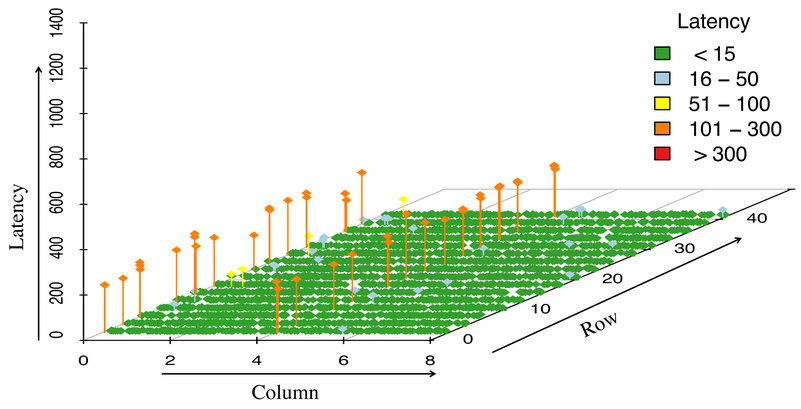
Approach

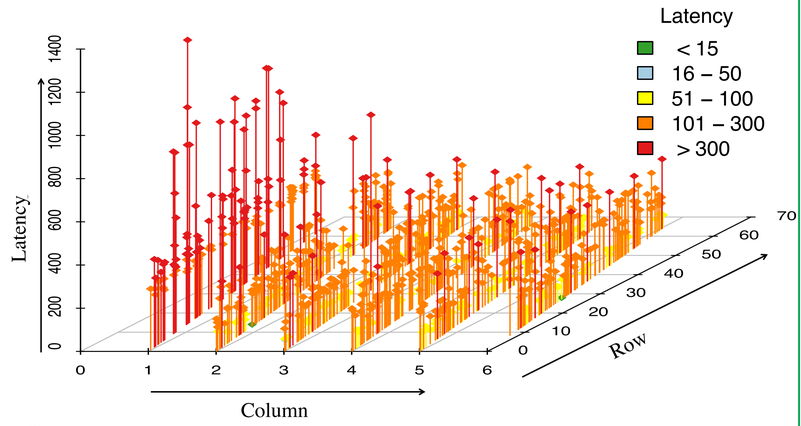
- Intel's new Precise Event Based Sampling (PEBS) -- hardware counters
- CPU generates a PEBS sample after randomly sampling "n" load instructions
- Matching information generated from Pin tool and PEBS samples enables us to identify the source of performance bottleneck



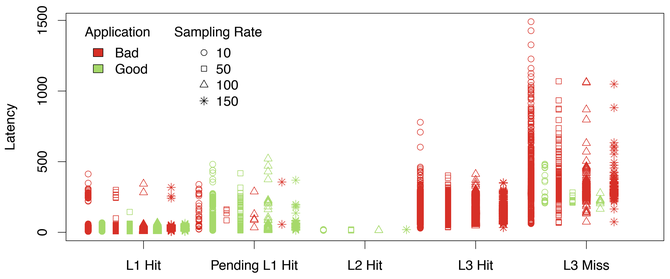
Preliminary Results







Sampling Rate	10		50		100		150	
Application	Good	Bad	Good	Bad	Good	Bad	Good	Bad
L1 Hit	99.748%	91.353%	99.739%	91.020%	99.820%	91.434%	99.802%	91.893%
Pending L1 Hit	0.216%	0.081%	0.225%	0.054%	0.135%	0.018%	0.0126%	0.072%
L2 Hit	0.027%	0%	0%	0%	0.036%	0%	0.018%	0%
L3 Hit	0%	4.058%	0%	5.992%	0%	5.983%	0%	5.453%
L3 Miss	0.009%	4.508%	0.036%	2.933%	0.009%	2.564%	0.054%	2.582%



Frequency of different events at different sampling rates

Future Plan

- Analyze real-world applications
- Pin-point specific variables causing performance bottleneck

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract LLNL-POST-569614.