Defensive Loop Tiling for Multi-core Processors

Bin Bao(包斌), Xiaoya Xiang(向晓娅)
University of Rochester

Presented by: Chen Ding(丁晨)
New Challenges in Multi-core Era

- Some levels of caches are shared by cores
- Different programs can compete for cache
- The available cache size is a run-time information
- A wrong selection of tile sizes can forfeit the benefit of tiling
Reuse Distance Analysis

- Cache configuration
  - L1: 32KB, private
  - L2: 256KB, private
  - L3: 8MB, shared
- Reuse distance

<table>
<thead>
<tr>
<th>Loop level</th>
<th>Reuse data</th>
<th>[256,2048,8]</th>
<th>[2048,2048,128]</th>
<th>[2048,128,128]</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>C[i][j]</td>
<td>32KB</td>
<td>32KB</td>
<td>2KB</td>
</tr>
<tr>
<td>i</td>
<td>B[k][j]</td>
<td>144KB</td>
<td>2065KB</td>
<td>130KB</td>
</tr>
<tr>
<td>kk</td>
<td>C[i][j]</td>
<td>4240KB</td>
<td>36MB</td>
<td>4224KB</td>
</tr>
</tbody>
</table>
- Matrix multiplication co-run with up to 3 STREAM benchmarks
- Search all the power of 2 tiling factors, \([TI,TJ,TK]\)

![Performance Comparison](image)

- **Alone**
  - [256,2048,8]
  - [2048,2048,128]
  - [2048,128,128]

- **1 STREAM**
  - [256,2048,8]
  - [2048,2048,128]
  - [2048,128,128]

- **2 STREAM**
  - [256,2048,8]
  - [2048,2048,128]
  - [2048,128,128]

- **3 STREAM**
  - [256,2048,8]
  - [2048,2048,128]
  - [2048,128,128]

**Execution time (seconds)**: 0, 2, 4, 6, 8, 10, 12, 17.7
Discussion

- Model cache interference as a function of tile size
- Balance the performance for different co-run environments
- Distinguish inclusive cache and non-inclusive cache
- Combine with prefetching effect