Automated Parallelization of Non-Uniform Convolutions on Chip Multiprocessors

Yuanrui Zhang, Mahmut Kandemir

Nikos Pitsianis, Xiaobai Sun
Computer Science, Duke University.

HPEC-2009, Sept 22-23, MIT Lincoln Laboratory
Non-Uniform FFT Local Convolution

\[ v(T) := F(T, \tilde{S}) \, u(\tilde{S}) \]

\[ c(T) \odot v(T) \approx F(T, S) \, C(S, \tilde{S}) \, u(\tilde{S}) \]

- While FFTs are well implemented by FFTW, we concentrate on accelerating the parallel convolution step on multicores.
- Geometric tiling can enhance data locality and reuse for the non-uniform local convolution, a matrix-vector product with an irregular and sparse matrix.
- Multicores have different on-chip memory hierarchies and characteristics.

HPEC-2009, Sept 22-23, MIT Lincoln Laboratory
Hierarchical tiling according to memory hierarchy and sizes

Neighborhood tile traversing order

Block distribution

Try to take care of data sharing and locality at all levels of cache

HPEC-2009, Sept 22-23, MIT Lincoln Laboratory
Preliminary Evaluation

Input size 6K×6K, \(\alpha=1.5\), \(W=10\), \(L=50\)

- Harpertown
- Nehalem
- Dunnington

<table>
<thead>
<tr>
<th>Machine</th>
<th>Execution Time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dunnington</td>
<td>600</td>
</tr>
<tr>
<td>Nehalem</td>
<td>500</td>
</tr>
<tr>
<td>Harpertown</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Machine</th>
<th>Normalized Execution Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dunnington</td>
<td>1</td>
</tr>
<tr>
<td>Nehalem</td>
<td>1</td>
</tr>
<tr>
<td>Harpertown</td>
<td>1</td>
</tr>
</tbody>
</table>

HPEC-2009, Sept 22-23, MIT Lincoln Laboratory