Automated Parallelization of Non-Uniform Convolutions on Chip Multiprocessors

Yuanrui Zhang, Mahmut Kandemir

Computer Science and Engineering, The Pennsylvania State University.

Nikos Pitsianis, Xiaobai Sun

Computer Science, Duke University.

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

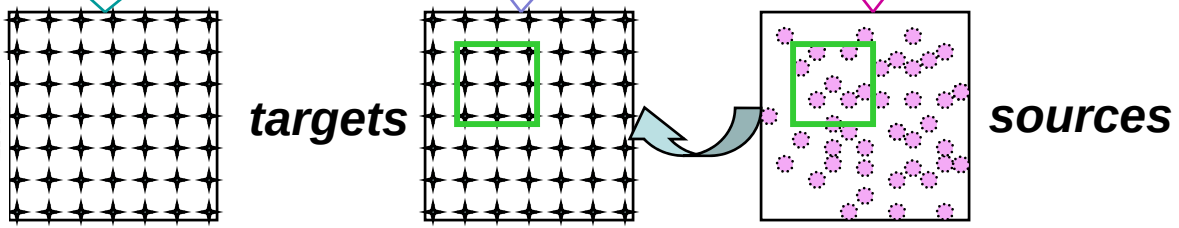
Non-Uniform FFT Local Convolution

$$\mathbf{v}(T) := \mathbf{F}(T, \tilde{S})\, \mathbf{u}(\tilde{S})$$

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

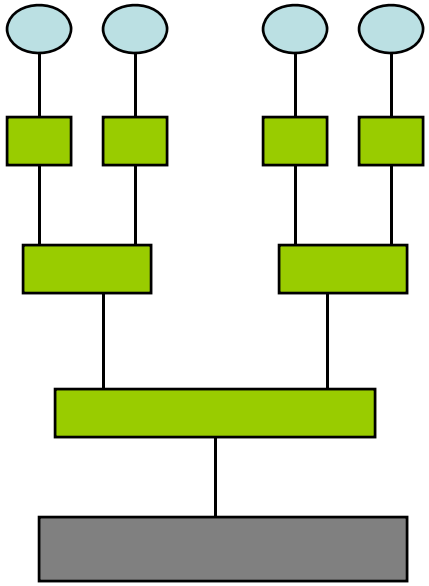
Local Support

$$\sum_{\|s_j - t_i\|_\infty < W} h(t_i - s_j)\, q(s_j)$$



- 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.

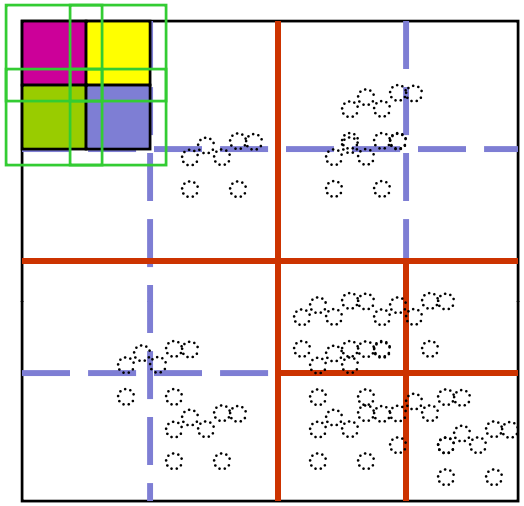
Architecture Aware Hierarchical Geometric Tiling and Parallel Scheduling

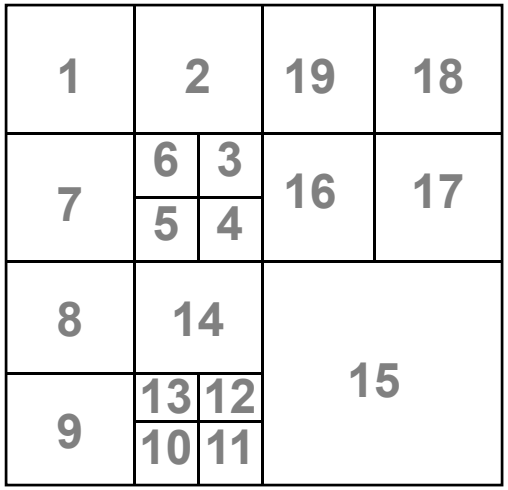


On-chip memory abstraction

- 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





Preliminary Evaluation