Problem Formulation

Current Data movement Strategy:
- Oblivious to the execution pattern of neighboring threads

Parallelism Vs Locality
- Accessing closed page is expensive both for performance and power
- Serialization of concurrent threads in memory banks

Aiming toward Exascale
- Stencil computation is at the heart of the simulation of several physical phenomena that is of scientific interest such as the Data Analysis Proxy App used in the Center for Exascale Simulation of Advanced Reactors.

Group Locality

Threads collaborate to improve locality using static information gained from the compiler.

The Fine Grain Execution Runtime

- Dependencies and tasks (i.e. tiles) are represented by bit masks.
- Dependency bit mask are AND’ed to get task mask.
- Inner tiles run in parallel as dependencies are satisfied.
- Take advantage of outer tile locality.
- Hierarchical

An Efficient Tiling Technique

The Jagged Tiling Technique
- Designed to improve intratile parallelism.
- Uses Polyhedral formulation to create the different tiling levels

- Two flavors:
  - Jagged Tiling (pipeline start codes) & Jagged Diamond Tiling

The Data Restructuring Framework

- Mem. access latency dependent on access pattern.
- Access pattern = Access Function + Mem. Mapping
- Based on Reuse of Data
- Mapping of one address space to the other using invertible transformation matrix.

Initial Results

- Increase performance by better utilization of resources
- Overall better energy profiles due to reduction in time and a negligible changes in power consumption
- Future work includes more in-depth characterization & memory structure mapping

Conclusions

Aiming toward Exascale
- Stencil computation is at the heart of the simulation of several physical phenomena that is of scientific interest such as the Data Analysis Proxy App used in the Center for Exascale Simulation of Advanced Reactors.

Authors: Sunil Shrestha1, Joseph Manzano2, Andres Marquez2, John Feo2 and Guang R. Gao1

1University of Delaware 2Pacific Northwest National Laboratory

This research was supported in part by DOE ASCR XStack program under Awards DE-SC0008716, DE-SC0008717