Adaptive Mesh Refinement under the Concurrent Collections Programming Model

Robert Pavel[1,2], Robert Bird[1,3], Pascal Grosset[1,4], Ken Czuprynski[1,5], Andrew Reisner[1,6], Erin Carrier[1,6], Christoph Junghans[1], Benjamin Bergen[1], and Allen McPherson[1]

[1]Los Alamos National Laboratory

[2]University of Delaware

[3]University of Warwick

[4]University of Utah

[5]University of Iowa

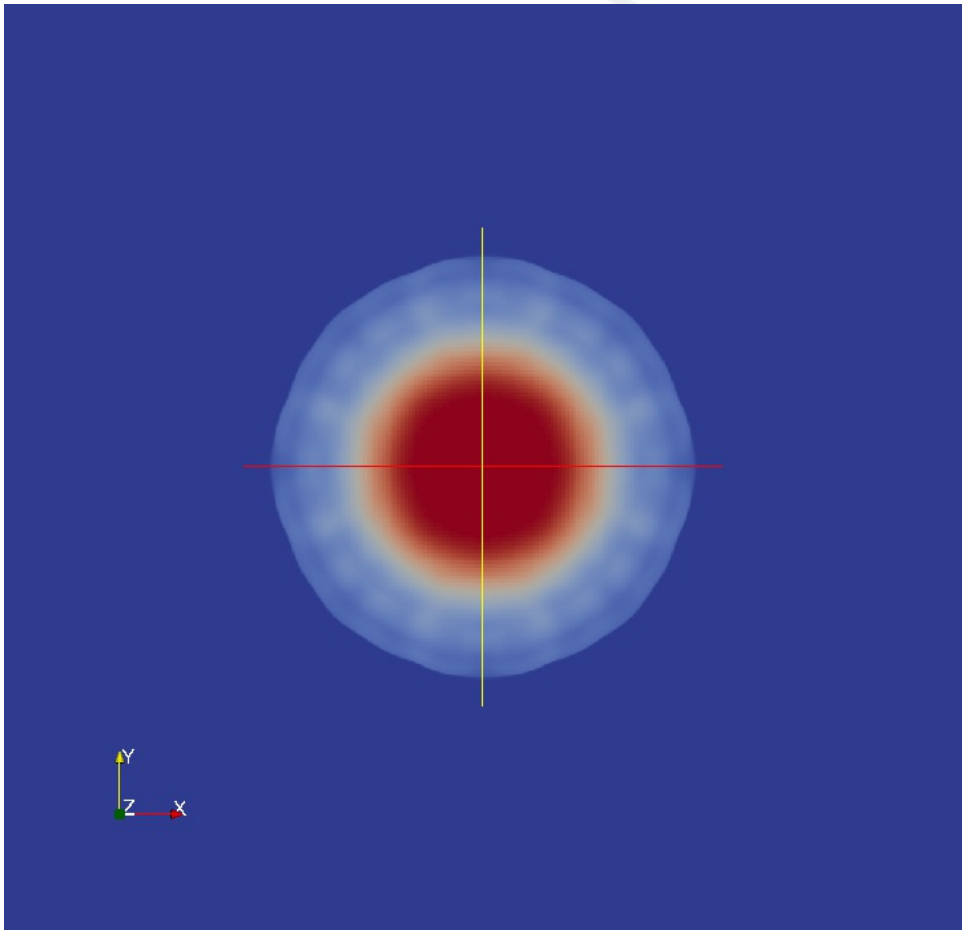
[6]University of Illinois at UrbanaChampaign

September 16, 2014

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Introduction

- Developed a scheme for Adaptive Mesh Refinement of a Hydrodynamics Simulation
  - The study of liquids in motion
  - Representative of several applications and interests in scientific computing
  - Tendency for very large scale simulations of non-uniform systems
- Worked with domain experts to develop general library
- Implemented in Intel's Concurrent Collections [1, 2]



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

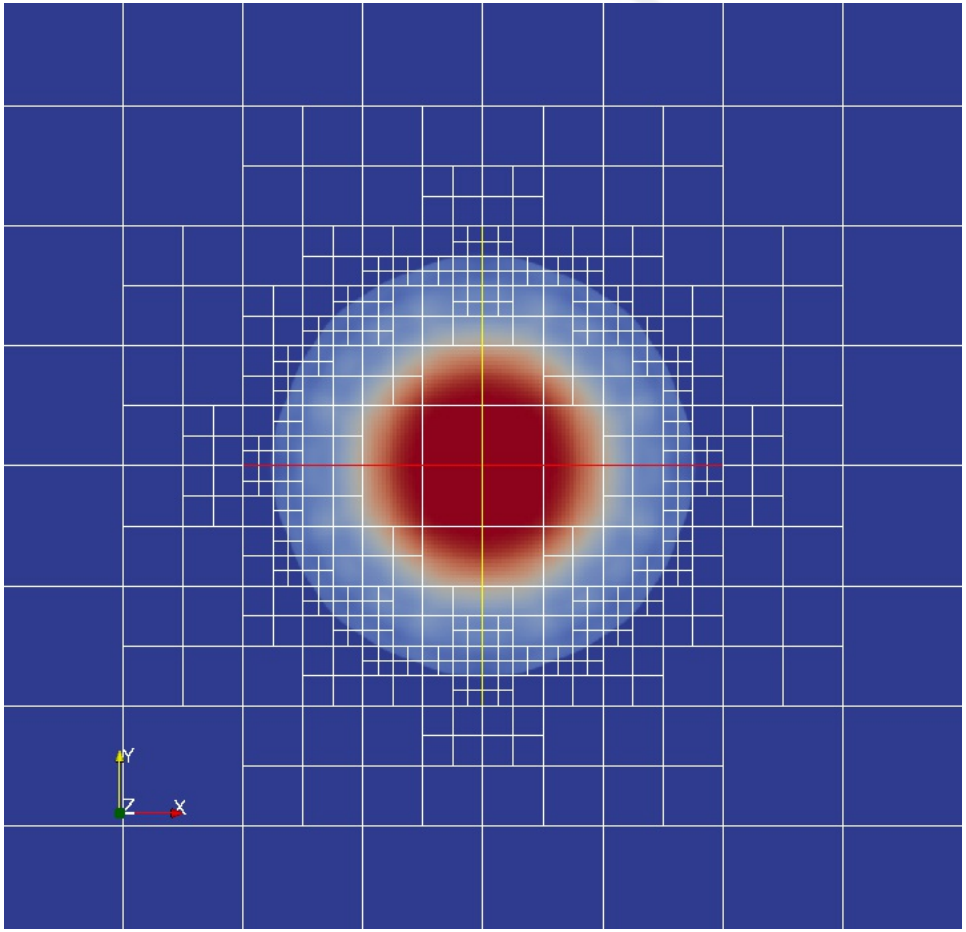
Adaptive Mesh Refinement

- Adaptive Mesh Refinement is a scheme for varying the accuracy of the solution based on the simulation state [3]
- Any given calculation is only as accurate as needed
  - i.e. Dynamically adjust the granularity of a region during execution
  - Areas with sharper discontinuities are "refined" to provide greater precision
  - Smoother regions are "coarsened" because less precision is needed
- Traditionally performed at either the Cell or Patch level
  - A Cell is the finest grain element of the grid
  - A Patch is a collection of contiguous Cells with the same refinement level

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Tile-Based AMR [4]

- Grid is divided into Tiles of varying levels of refinement
- Tiles are rectangular collections of Cells of uniform refinement
- Every Tile contains the same number of Cells
- Decision to refine or coarsen made at the Tile level
- During Refinement and Coarsening, Tiles are split or merged
- Refinement level corresponds to physical dimensions of Cells in Tile
- Can be considered a subset of Patch-Based AMR with stricter rules regarding refinement



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Challenges of AMR

- Rapidly varying workload
  - Refinement and coarsening changes the number and/or weight of tasks
- Non-Uniform Grid
  - Need to reconcile differences between refinement levels
- Data Storage
  - Need to ensure location of all required data is known and available

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Co-Design

- Non-conventional design philosophy in which domain experts and computer scientists work in unison
- Traditionally domain experts develop simulations and computer scientists later optimize or re-implement for performance
- Under co-design, a multi-disciplinary approach is employed
- This work developed as part of the 2014 Los Alamos National Laboratory Co-Design Summer School
  - Multi-disciplinary team approached a problem of interest to the laboratory
  - Explored alternative solutions to problem

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

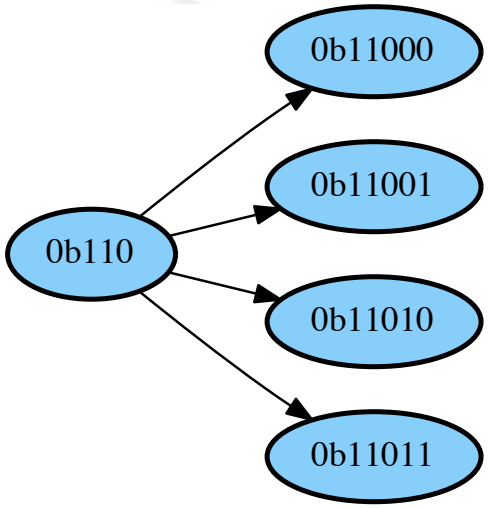
Our Approach

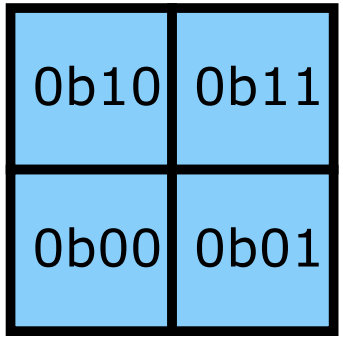
- Studied existing approaches
- Settled on use of modern runtime systems and programming models to simplify problem
  - Programming models that are better suited to the underlying science
  - Built in mechanisms for load balancing
- Worked with domain experts to make flexible, quad-tree based, AMR library with modular solvers
- Used library to implement application in multiple programming models and runtimes
  - (Intel) Concurrent Collections is one of the models we chose

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

High Level Details

- Hydrodynamics simulation of propagation of a shockwave through system
  - Finite Volume Method
  - Dimensional Splitting
  - Solve 2D Euler Equations with MUSCL-Hancock Scheme [5]
- Tile-based AMR
  - Structured via map-based Quadtree [6]
  - Morton Encoding (Z-Order) scheme used to provide Tile IDs [7]
  - Allows Quadtree to be represented by a map containing only leaves
  - Limit the number of different refinement levels to limit error
- Focus on Library Based, Functional, Approach
  - Clearly defined interfaces to all aspects
  - Modular approach to allow multiple solvers to be used
  - So as to facilitate integration with multiple programming models





This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

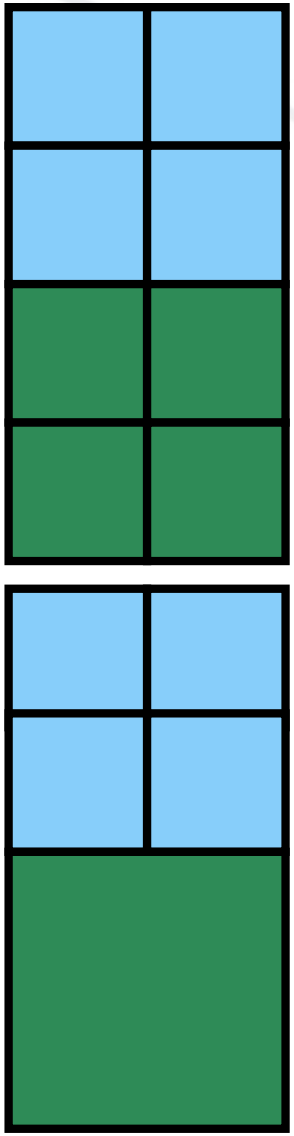
The Algorithm

- Determine length of timestep (dt) based on Courant–Friedrichs–Lewy (CFL) conditions [8]
  - Ensures wave does not propagate beyond bounds of tile in a single timestep.
- Update simulation in X-Dimension
  - Exchange Ghost Cells
  - Perform a sweep on all Cells in Tile
  - Apply Flux Corrections at discontinuities in refinement level
- Update simulation in Y-Dimension
- Apply Refinement Strategy to determine need for refinement or coarsening
- Refine and Coarsen as needed

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Refinement and Coarsening

- Overall grid requires 2:1 rule to be enforced
  - A given Cell/Tile may have up to two neighbors along any edge
  - Drastically simplifies implementation while providing bounds on error
- Refinement is high priority
  - One (parent) Tile Refines into four child Tiles (siblings)
  - If refinement violates 2:1, neighbors are also refined until 2:1 is not violated
- Coarsening is low priority
  - Four sibling Tiles Coarsen into one (parent)
  - May Coarsen only if all four are flagged to and 2:1 won't be violated



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

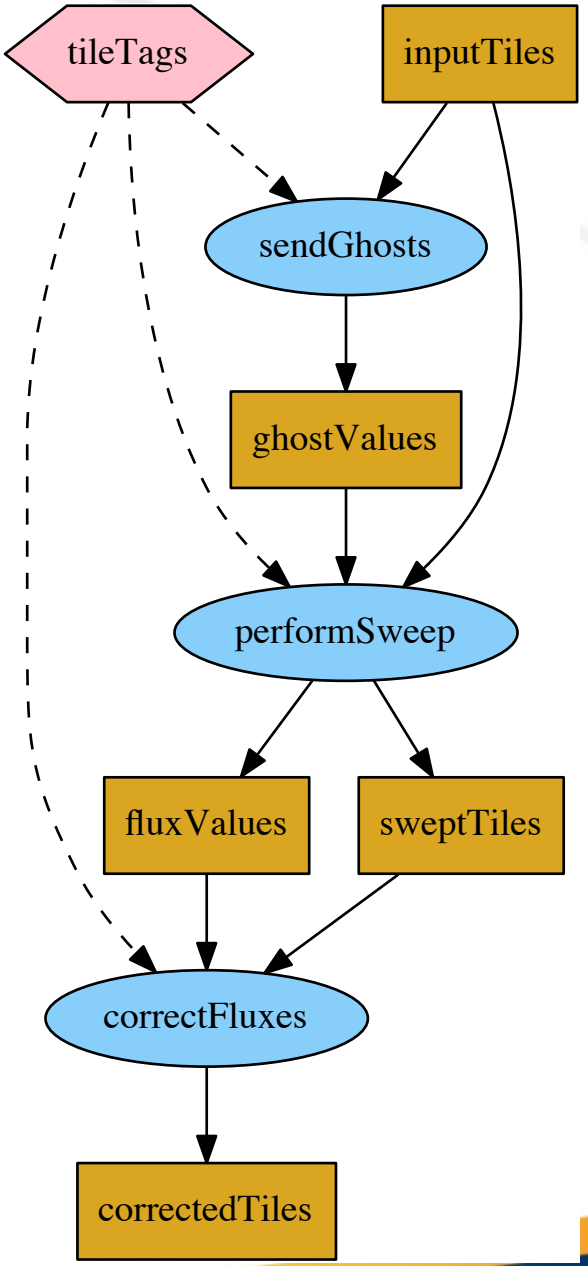
Implementation in CnC

- Generally straightforward
- Communication almost exclusively producer/consumer relationships
- Requires reduction on *dt*
- Item Collections a good fit
  - Underlying quadtree based on map data structure
- Write–once memory problematic for refinement

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

A Dimensional Sweep

- A given dimensional sweep is a stencil with very simple dependencies
  - Tile from previous Step
  - Ghost Cells from nearest neighbors in the dimension we are sweeping
  - Flux corrections from neighbors in dimension we are sweeping with finer refinement levels
- Very simple CnC implementation
  - Each Tile has a neighbor list
  - All dependencies easily expressed through producer/consumer relationships



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

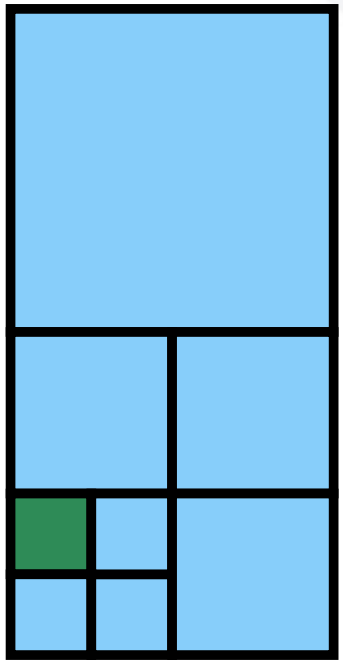
Apply Refinement Strategy

- Uses Cells in Tile to determine if Refinement or Coarsening is needed
  - Estimates error based on contents of Tile
  - If error is above a threshold, Refinement occurs
  - If error is below a threshold, Coarsening occurs
- Specifically, we utilize the scheme developed by Löhner
  - Makes need for ghost Cells less important
  - Lessens communication
- Also straightforward
  - Similar communication pattern to a Dimensional Sweep

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

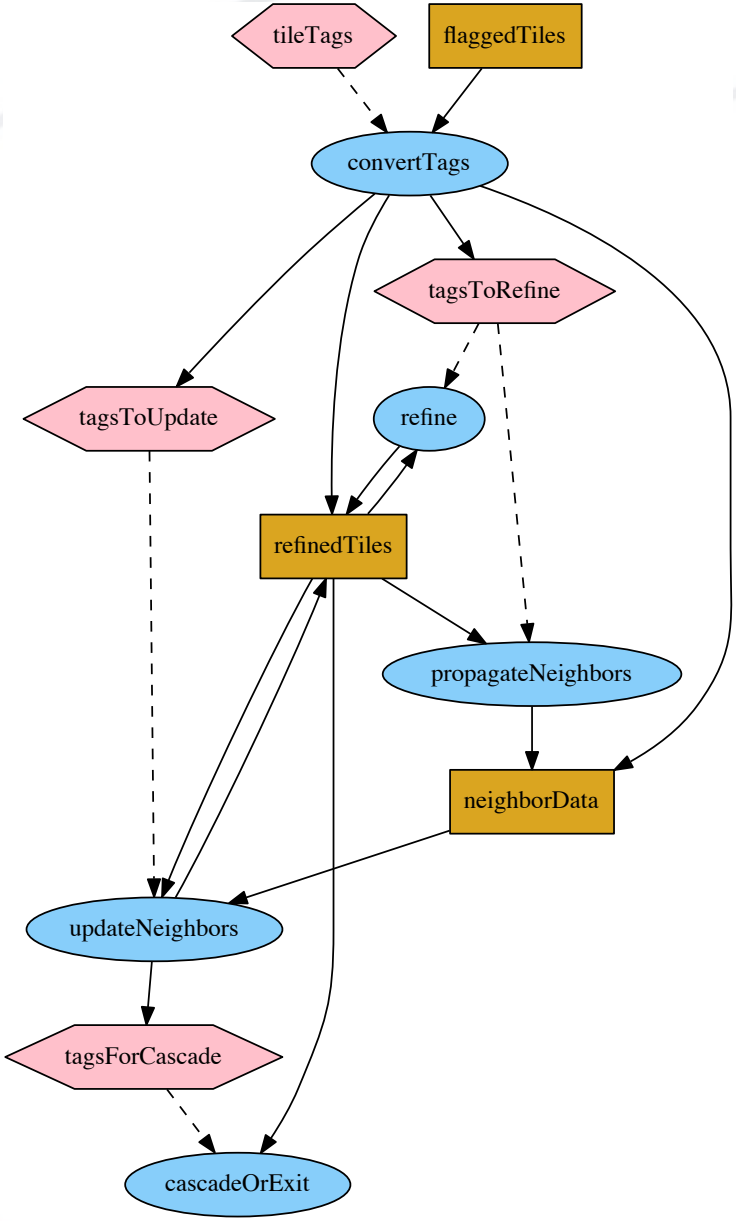
Refinement

- Refinement itself is complicated
- High priority of refinement leads to refinement cascades
  - Cascades can trigger additional cascades
- Items are write–once memory
- Need to ensure all Items are written to the correct location once
- Maximum number of cascades related to number of different refinement levels
- Every Tile steps through the refinement loop once for each refinement level
  - Ensures that tag of each Tile is known and available and that final result is in correct location
- Aside from that, fairly straightforward



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

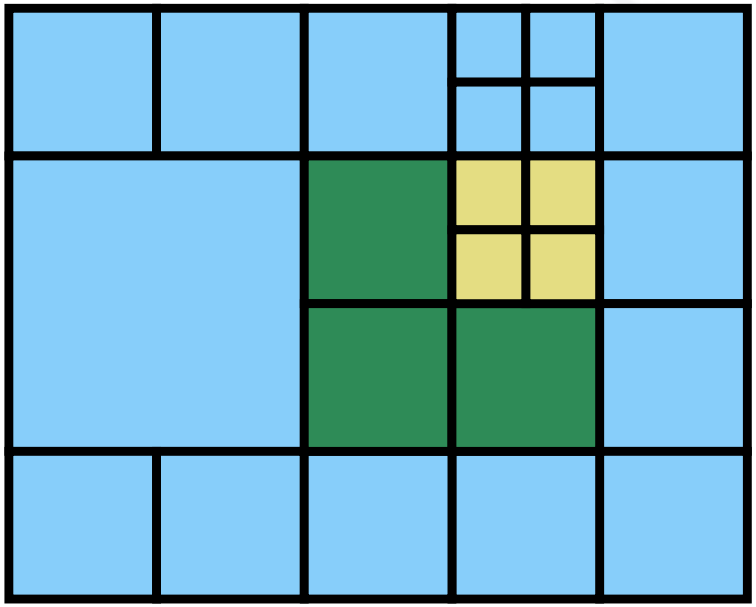
Refinement



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

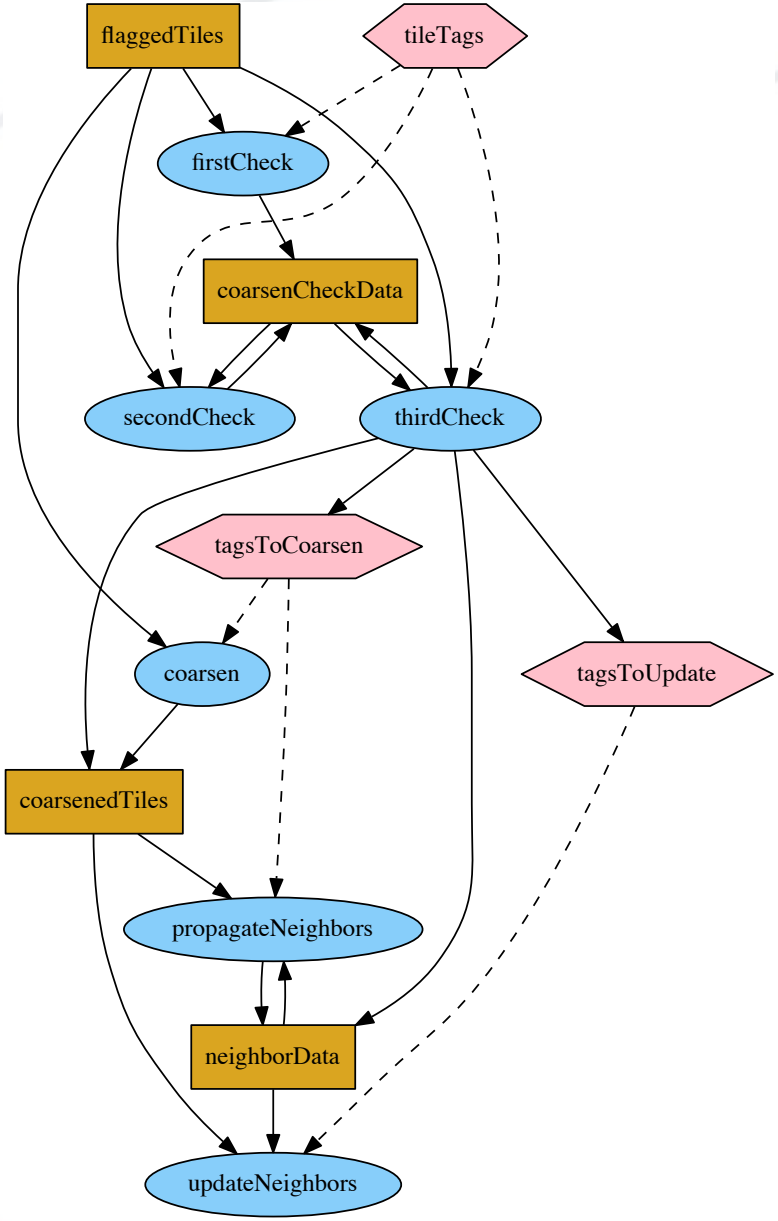
Coarsening

- Coarsening much simpler
- Coarsening is low priority
  - All four siblings must be flagged to coarsen
  - Coarsening must not violate 2:1 rule
- Each sibling locally determines if it can coarsen
  - Are all non-sibling neighbors as refined or coarser?
  - Do all sibling neighbors exist at the same refinement level?
- This information is propagated to siblings
- If all four are able, coarsening occurs



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Coarsening



This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Reduction on *dt*

- To satisfy CFL conditions, each Tile locally computes its desired *dt* for the next time step
- A reduction is then needed to find the global *dt*
- A CnC reduction requires knowledge of the number of participants
- The number of tiles varies during execution beause of adaptive mesh refinement
- This is solved through multiple reductions
  - One reduction to determine the number of Tiles after refinement and coarsening
  - One reduction to obtain minimum *dt*
- Only need to wait on result of *dt* reduction
  - *dt* reduction needs to wait on result of tile reduction
  - Performance depends heavily on how much of a reduction can occur before the number of participants is known

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

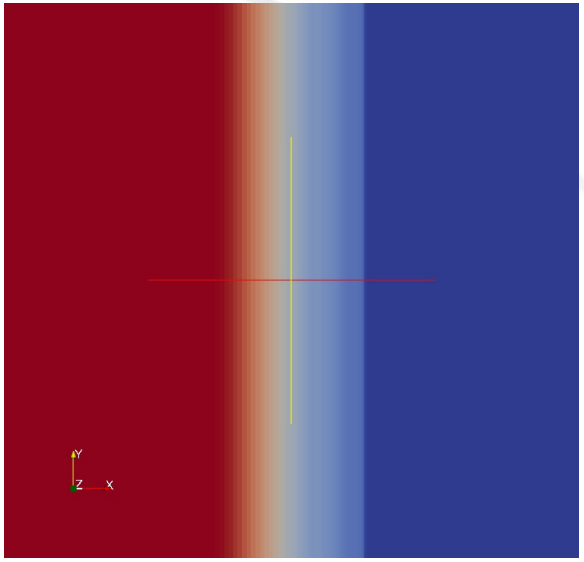
Tuning and Optimizations

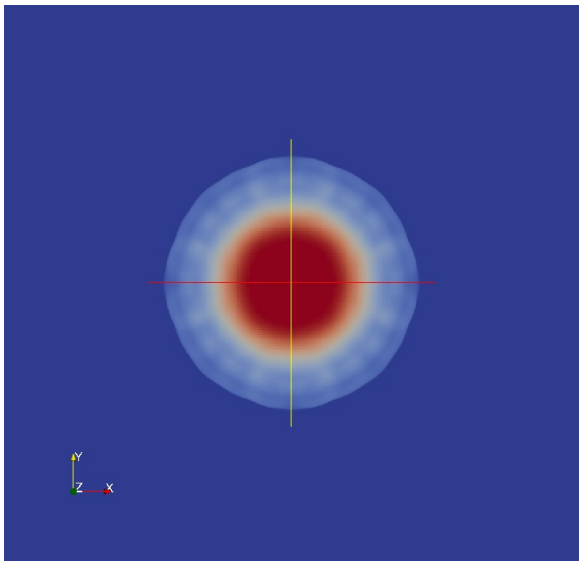
- Still ongoing
- Step scheduling (Step Tuners) an open problem
  - Simple modulo based scheduling insufficient due to encoding scheme
  - Would prefer to have locality, at least, during Coarsening
  - Traditional solutions (ParMETIS [9], PT-SCOTCH [10], etc.) are heavyweight
- Garbage Collection (Item Tuners)
  - Mostly done
  - Need to adjust Refinement and Coarsening branches to avoid accessing Items a varying number of times
- Make tasks coarser–grained
  - Operate on a "branch" instead of a "leaf"
  - Guarantees locality during Coarsening
  - May lead to more complicated load balancing during Refinement

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Preliminary Results

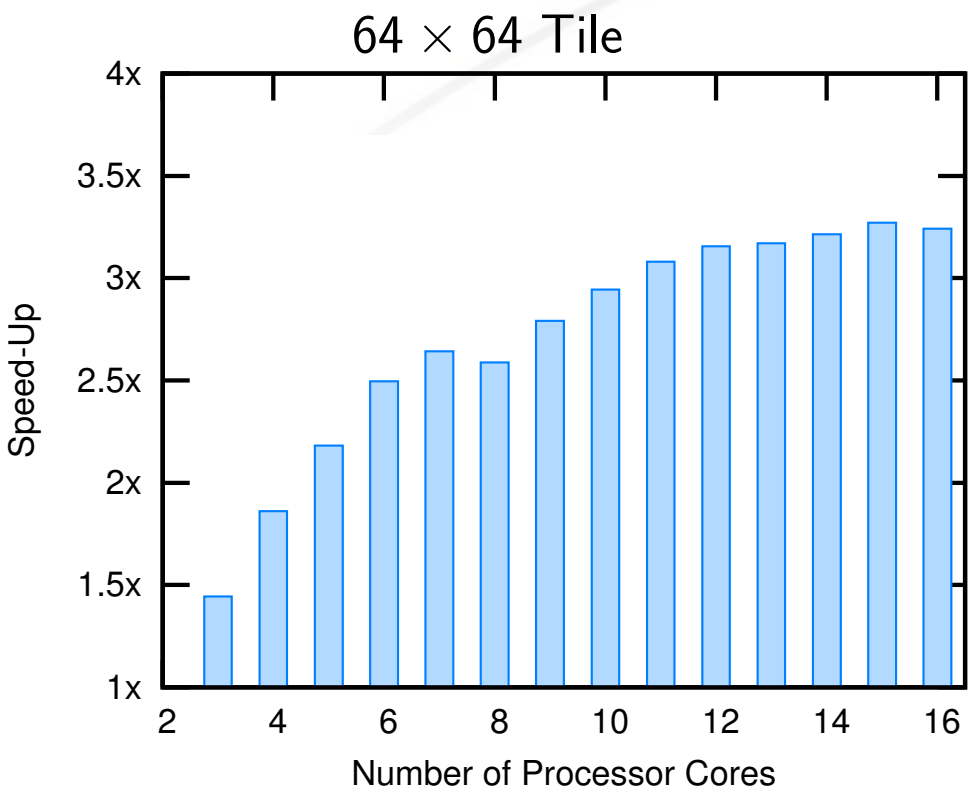
- Very early results
- Focused on shared memory initially
- Considered two Tile sizes and two physical problems
  - Sod Shock
  - Centralized Discontinuity
- Limited to four levels of refinement for memory purposes
- Tests run on single node of Los Alamos National Laboratory's Darwin cluster
  - 2 Socket Intel Xeon E5-2650L. 1.8 GHz, 16 physical cores, 128 GB memory

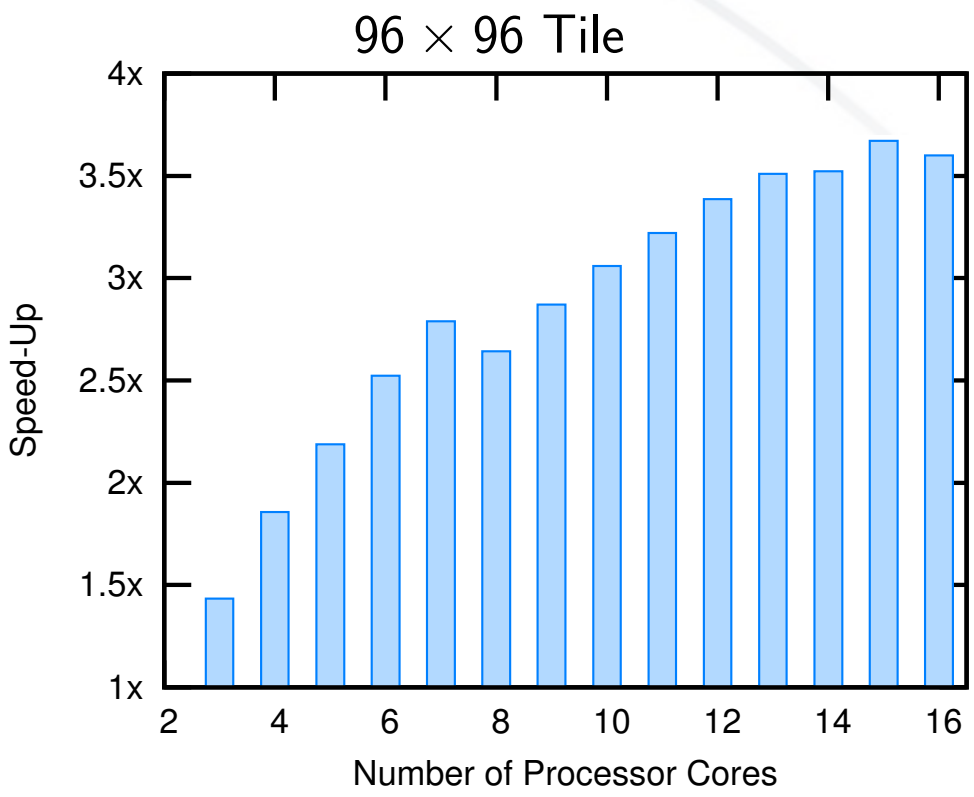




This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

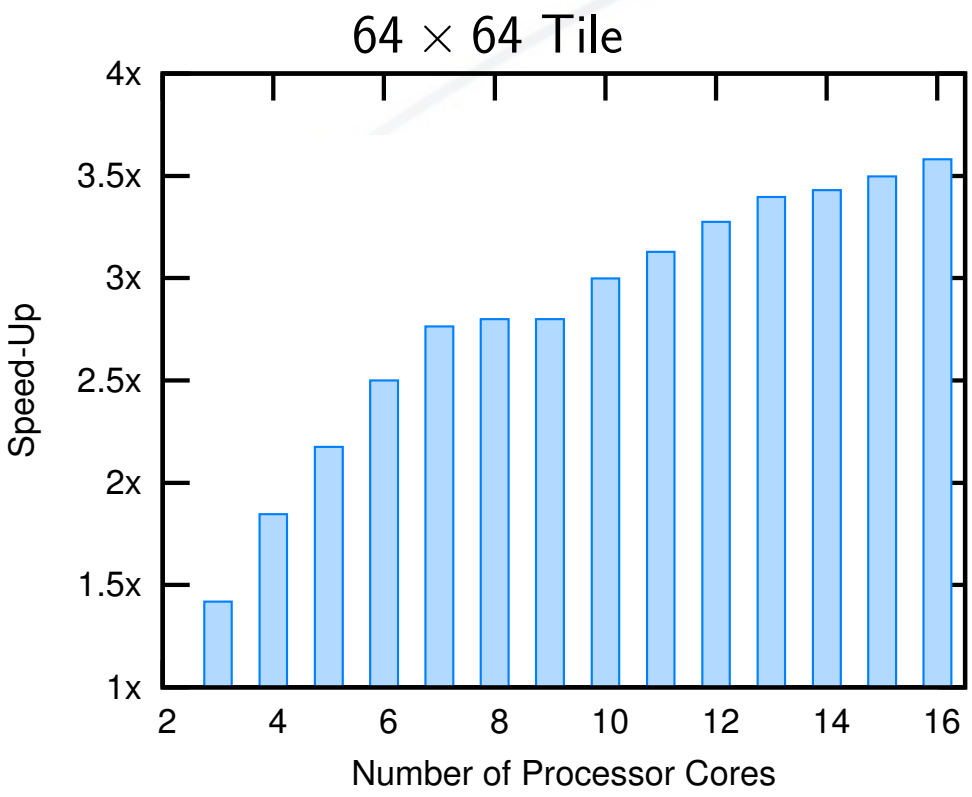
Sod Shock

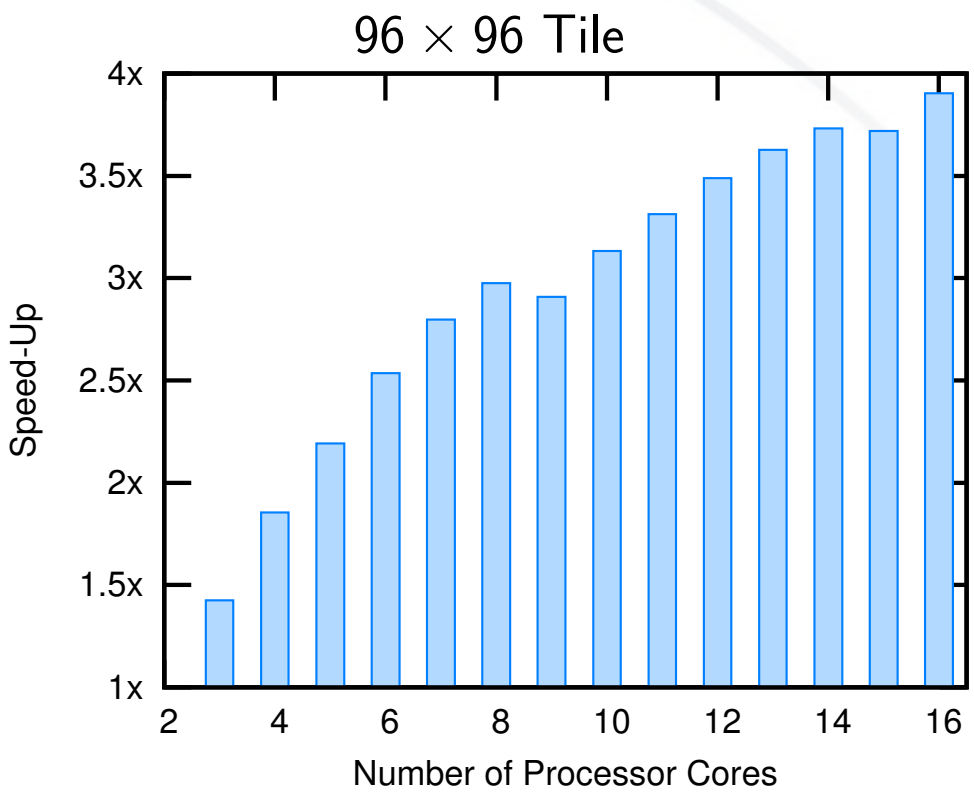




This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Centralized Discontinuity





This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Suitability of CnC

For this application, CnC was an excellent fit

- Majority of work was a finite volume problem with simple dependencies
- Underlying Data Strucure was built around a Map
  - CnC Item Collections were a good fit
  - Tag–space Extensions used to handle iterations
- Write–Once Memory was problematic for Refinement
  - Workaround includes unnecessary computation and communication
- Built-in debugging and tracing tools invaluable in development of implementation

For other applications, CnC may cause issues

- Not all data structures can be easily represented with a Write–Once map
- Write–Once Memory makes unbounded loops problematic
  - Many applications run until convergence criteria are met
  - Unclear of efficient way to check those criteria (periodic Reductions?)

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Acknowledgments

- This work was supported by the Los Alamos Information Science and Technology (IS&T) Co-Design Summer School and the U.S. Department of Energy (DOE) Office of Advanced Scientific Computing Research (ASCR)
- Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Los Alamos National Security, LLC for the National Nuclear Security Administration (NNSA) of the U.S. DOE under contract DE-AC52-06NA25396
- The authors would also like to acknowledge the Institutional Computing Program at LANL for use of their HPC cluster resources. Additional simulations were performed on the CCS-7 cluster Darwin.

References I

- K. Knobe, "Ease of use with concurrent collections (CnC)," *Hot Topics in Parallelism*, 2009.
- M. G. Burke, K. Knobe, R. Newton, and V. Sarkar, "Concurrent collections programming model," in *Encyclopedia of Parallel Computing*. Springer, 2011, pp. 364–371.
- M. J. Berger and J. Oliger, "Adaptive mesh refinement for hyperbolic partial differential equations," *Journal of Computational Physics*, vol. 53, no. 3, pp. 484–512, Mar. 1984. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0021999184900731
- Q. F. Stout, D. L. De Zeeuw, T. I. Gombosi, C. P. T. Groth, H. G. Marshall, and K. G. Powell, "Adaptive blocks: A high performance data structure," in *Proceedings of the 1997 ACM/IEEE Conference on Supercomputing*, ser. SC '97. New York, NY, USA: ACM, 1997, pp. 1–10. [Online]. Available: http://doi.acm.org/10.1145/509593.509650
- *Riemann Solvers and Numerical Methods for Fluid Dynamics - A Practical Introduction*. [Online]. Available: http://www.springer.com/materials/mechanics/book/978-3-540-25202-3

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

References II

- H. Samet, "The quadtree and related hierarchical data structures," *ACM Comput. Surv.*, vol. 16, no. 2, pp. 187–260, Jun. 1984. [Online]. Available: http://doi.acm.org/10.1145/356924.356930
- M. S. Warren and J. K. Salmon, "A parallel hashed oct-tree n-body algorithm," in *Proceedings of the 1993 ACM/IEEE Conference on Supercomputing*, ser. Supercomputing '93. New York, NY, USA: ACM, 1993, pp. 12–21. [Online]. Available: http://doi.acm.org/10.1145/169627.169640
- P. Lax and B. Wendroff, "Systems of conservation laws," *Communications on Pure and Applied Mathematics*, vol. 13, no. 2, pp. 217–237, May 1960. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/cpa.3160130205/abstract
- G. Karypis, K. Schloegel, and V. Kumar, "Parmetis: Parallel graph partitioning and sparse matrix ordering library," *Version 1.0, Dept. of Computer Science, University of Minnesota*, 1997.
- C. Chevalier and F. Pellegrini, "PT-scotch: A tool for efficient parallel graph ordering," *Parallel Computing*, vol. 34, no. 68, pp. 318–331, Jul. 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167819107001342

This work approved for unlimited public release as per LA-UR-14-26324
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA