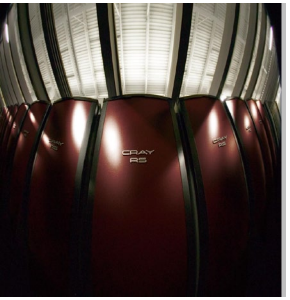


Lazy, Minimal, Eventually Consistent IO with Stitch

Jay Lofstead, John Mitchell

PDSW-DISCS WIP

November 12, 2018

SAND2018-12554 C

Exceptional service in the national interest

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Problem Space

- Kinetic Monte Carlo simulations
  - Welding, additive manufacturing
- Finite Element
  - Pressure wave from striking a rod
- Others certainly possible

- Computation is localized to a small part of the compute domain. Why do compute and IO traditionally when we can be MUCH more efficient?

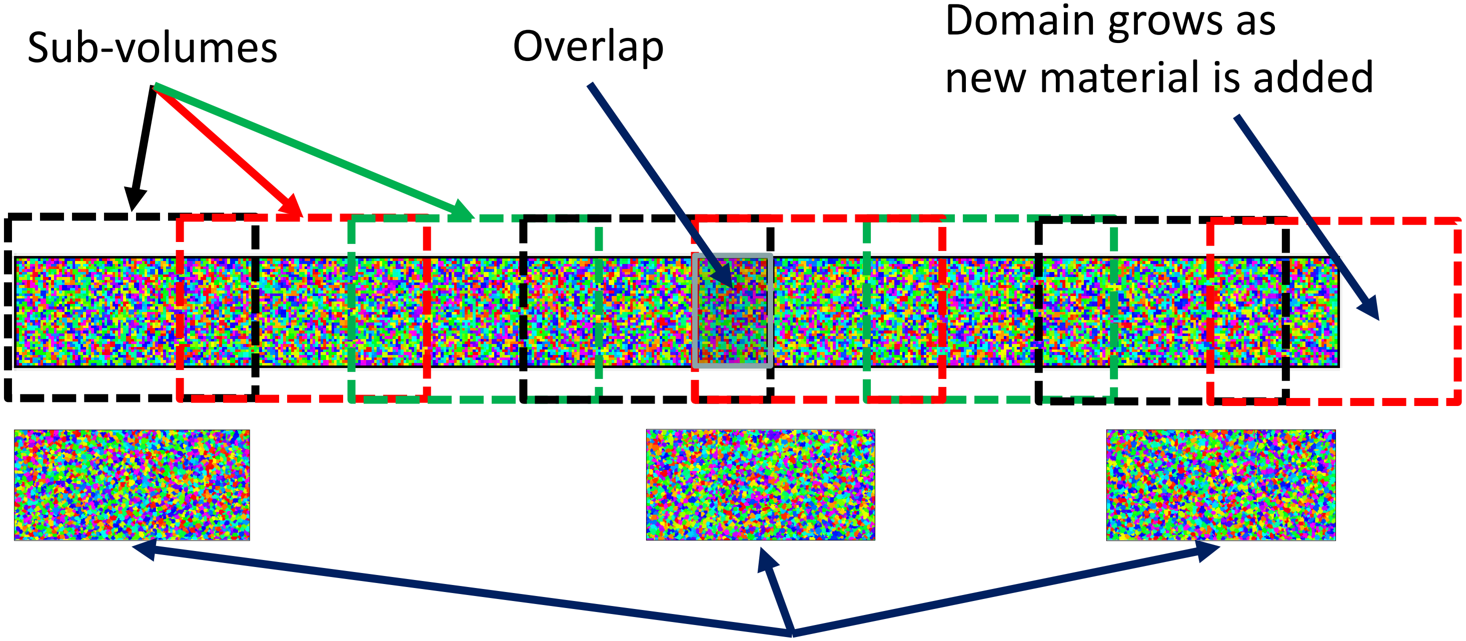
Approach

- Lazy
  - Only track what has been seen so far (i.e., we don't care about the size of the simulation domain)
- Minimal
  - Only write was has changed since last output
- Eventually Consistent
  - Rely on the output to eventually "make sense"

- Reading specifies an arbitrary region and a time; Stitch assembles ('stitches') the region state together from various pieces using the newest for every point

Illustrative Example (SPPARKS)

Grain growth across a large domain is simulated using a series of smaller overlapping sub-volumes.



Post-process, visualize and analyze on arbitrary sub-volumes and arbitrary times

Benefits and Challenges

- Move from 1024 process to 16
- Move data size to 1/64th size per output
- Wall clock time the same or slightly smaller (less output time)

- LOTS more features we can talk about offline
- Open Source (LGPL) in the final moments of completion (look at my github [Jay Lofstead (gflofst)])
- Full paper coming H1 2019