Reading an r-file material model in SW4 using a parallel file system

USGS workshop, Menlo Park, CA

N. Anders Petersson, Bjorn Sjogreen

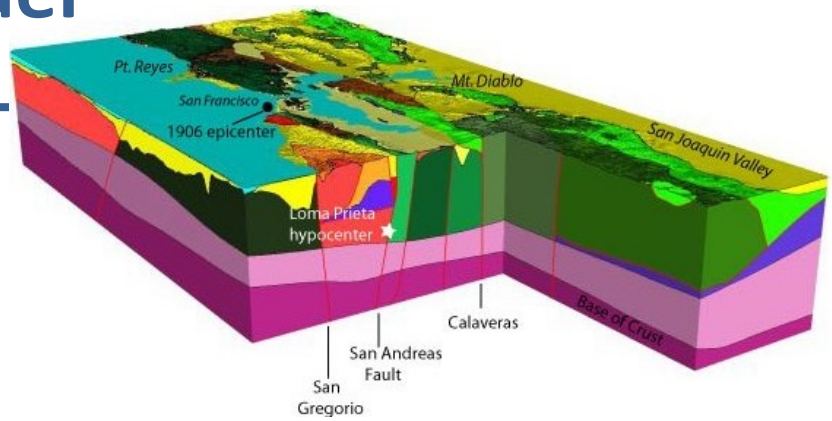
March 2018

LLNL-PRES-729414

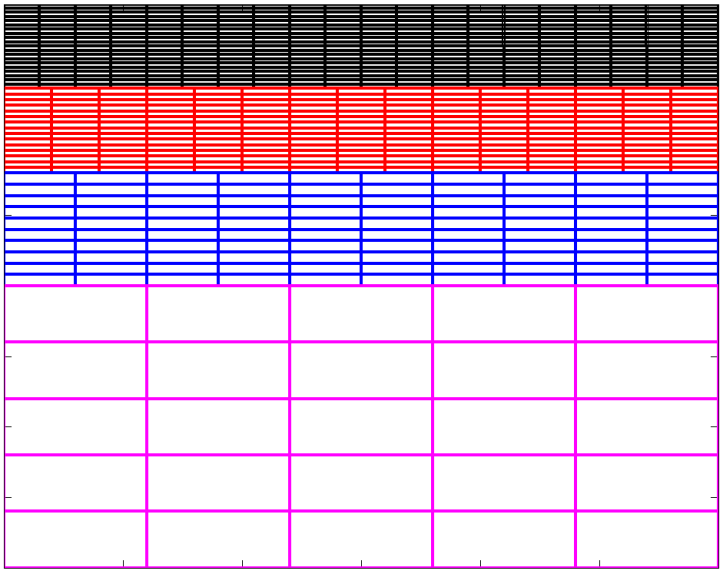
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

USGS Bay Area Material Model

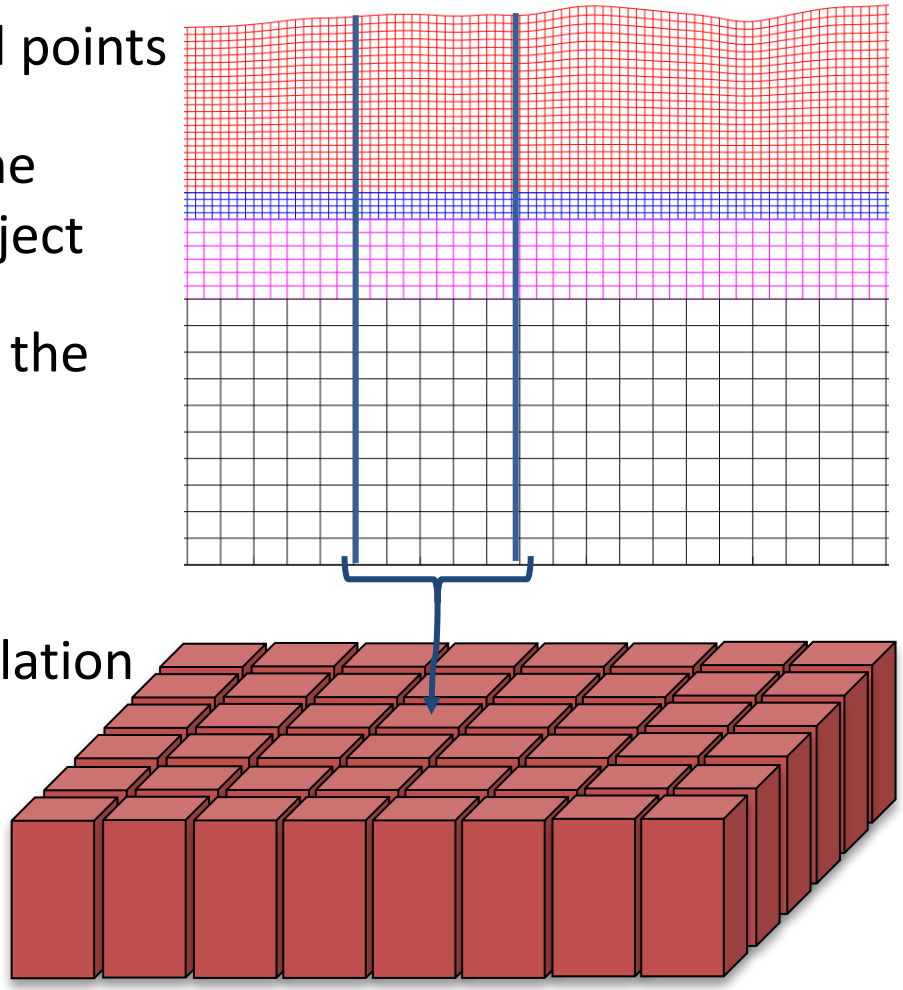


- Oct-tree variable resolution cell-centered data structure
- 8 Gbyte file (regional)
- Proj4 cartographic mapping (lon,lat,depth)-> (x,y,z)
- cencalvm: Point-wise query functions in C++
- R-file binary format developed in 2014
  - Byte-swapping for IBM BG/Q system at LLNL
  - Faster access; parallel file-system
- Header + rectangular blocks of vertex-centered data
- Same data as in the USGS e-tree model
- Read (Nc x Ni x Nj x Nk) blocks of data

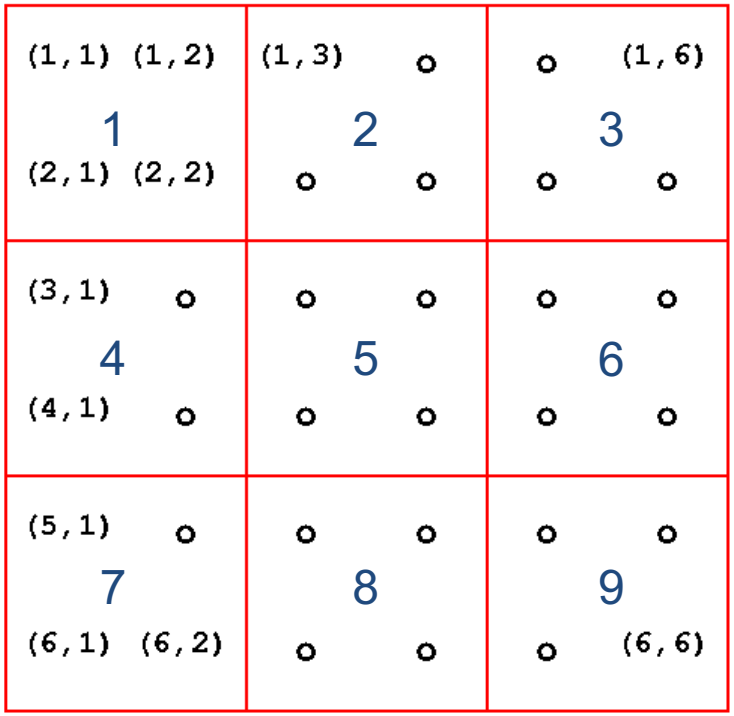


In SW4, only a subset of MPI-tasks read the r-file from the parallel file system

- Each MPI-task owns a “pencil” of grid points
- Each MPI-task requests a subset of the material model from a Parallel_IO object
- Only a subset of MPI-tasks read from the parallel system (8-128 readers)
- Material data is distributed by MPI
- Grid point values by tri-linear interpolation

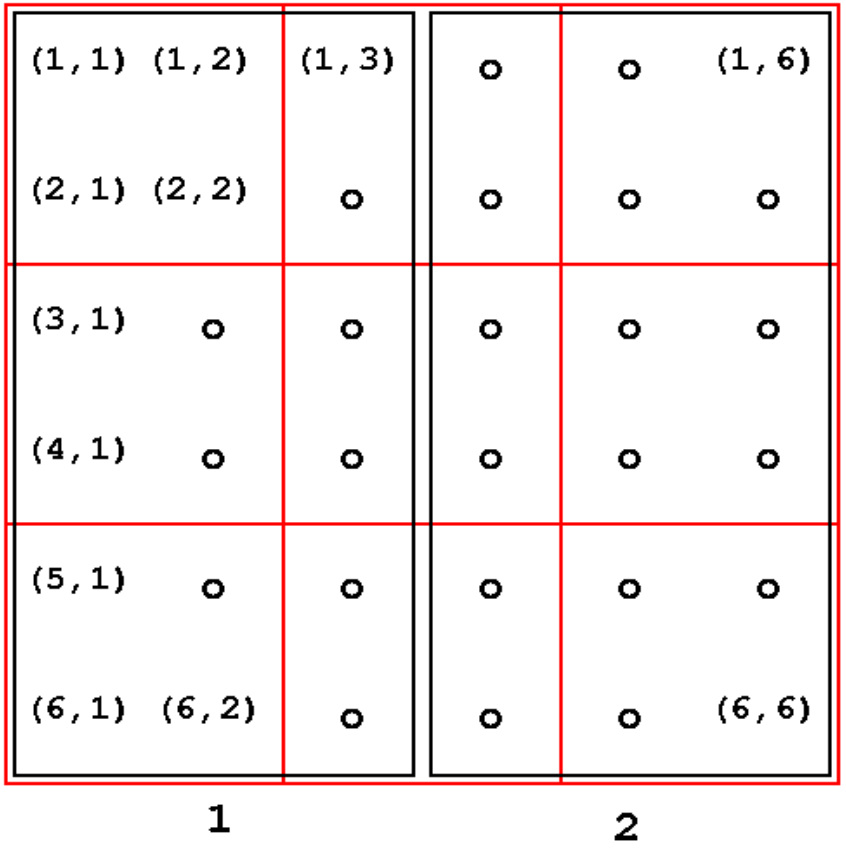


Example in 2D, 6x6 matrix on 9 MPI-tasks



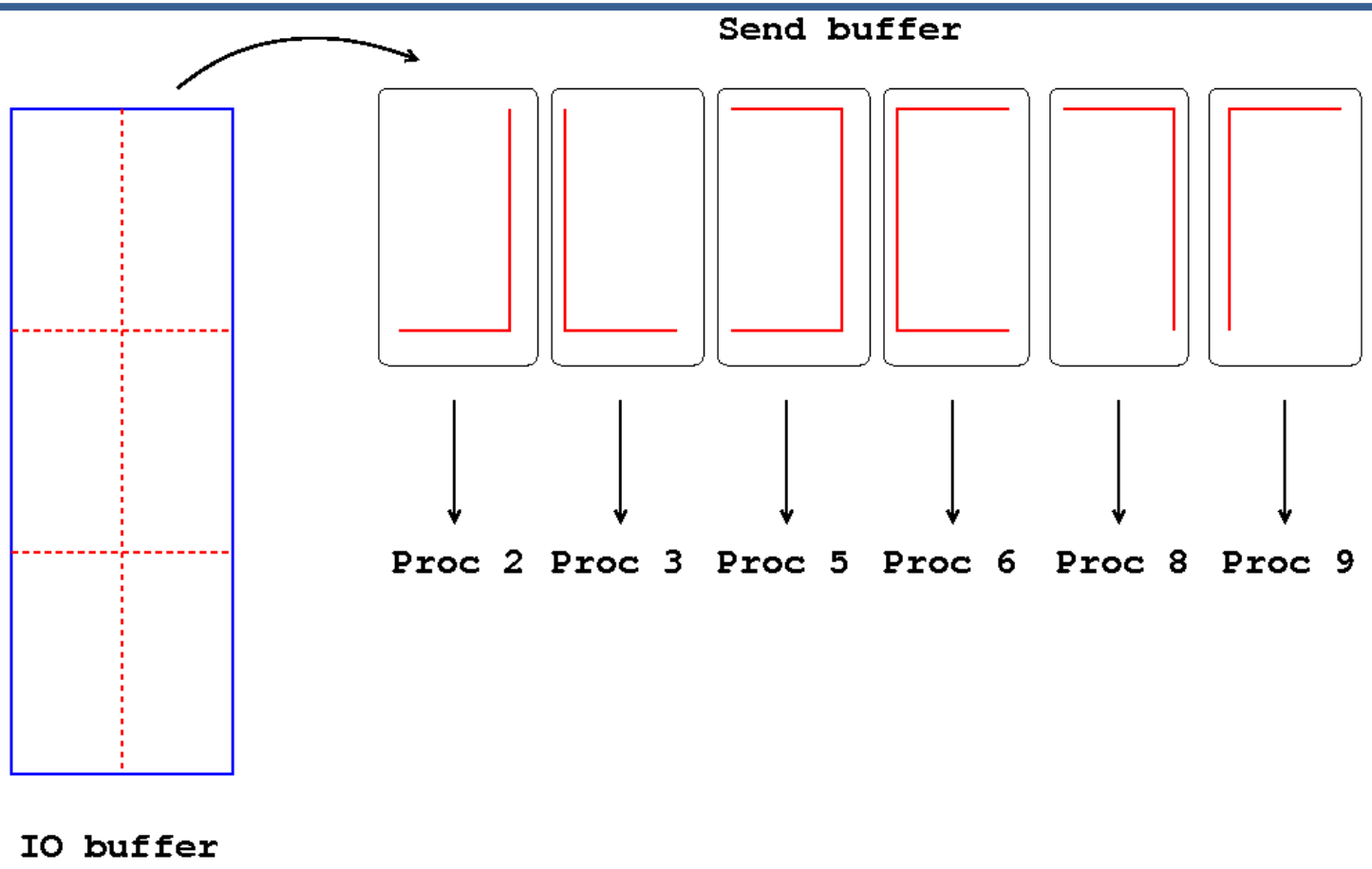
On disk in column order: (1,1), (2,1), (3,1)… (6,1), (1,2), (2,2), … (6,6).

Let 2 MPI-tasks be designated IO processors



Designated I/O processors

Each IO-processor reads data into a buffer and sends the requested subsets with MPI



I/O processor 2

Summary

- R-file format holds the same data as the USGS e-tree format
- Each MPI-task requests a subset of the material model
- 8-128 I/O processors read data from the parallel file system
- ~5min to read and initialize 60 billion grid point model on 2,048 nodes (131,072 MPI-tasks, 64 readers)
- Parallel IO routines originally developed for CFD applications:
  - B.Sjogreen, H.C.Yee, M.J.Djomehri, A.Lazanoff, and W.D.Henshaw (2010) "Parallel performance of ADPDIS3D - A high order multiblock overlapping grid solver for hypersonic turbulence", in "Parallel Computational Fluid Dynamics", R.Biswas ed., DEStech Publications.


Lawrence Livermore
National Laboratory