Large-Scale Parallel Performance of Algebraic Multigrid Preconditioners for Multiphysics Systems

Paul Lin, John Shadid, Ray Tuminaro
Sandia National Labs

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Abstract

The efficient computation of multiphysics simulations is challenging for several reasons: strong coupling of the multiphysics systems, high nonlinearity and a large spectrum of interacting length and time scales. This study evaluated the parallel scaling and performance of two multiphysics application codes for large-scale simulations. One application code simulates the drift-diffusion equations for semiconductor devices while the other simulates resistive MHD. A finite element method is used to discretize the system of PDEs on an unstructured mesh, then a fully-implicit Newton-Krylov solution method is employed. High fidelity solutions require the efficient and scalable solution of the large sparse linear systems. We considered the scaling of three different algebraic multigrid aggregation schemes for the drift-diffusion system, and found that the choice of aggregation scheme can have a significant impact on performance. Weak scaling studies for the drift-diffusion system have been performed up to 140,000 cores on both Blue Gene/P and Cray XE6 platforms.

Resistive MHD Model

(J. Shadid, R. Pawlowski, E. Cyr, L. Chacon)

Navier-Stokes + Electromagnetics

$$\rho\frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u}\cdot\nabla\mathbf{u}) - \nabla\cdot(\mathbf{T} + \boxed{\mathbf{T_M}}) - \rho\mathbf{g} = 0$$

$$\mathbf{T} = -(P + \frac{2}{3}\mu(\nabla\cdot\mathbf{u}))\mathbf{I} + \mu[\nabla\mathbf{u} + \nabla\mathbf{u}^T]$$

$$\mathbf{T_M} = \frac{1}{\mu_0}\mathbf{B}\otimes\mathbf{B} - \frac{1}{2\mu_0}\|\mathbf{B}\|^2\mathbf{I}$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0$$

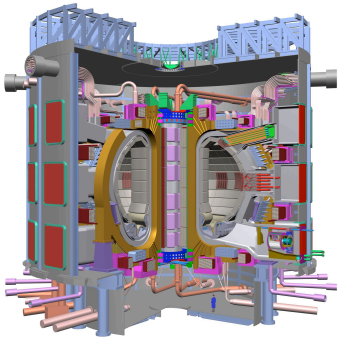
$$\rho C_p\left[\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T\right] + \nabla\cdot\mathbf{q} - \boxed{\eta\|\mathbf{J}\|^2} = 0$$

$$\boxed{\frac{\partial \mathbf{B}}{\partial t} - \nabla\times(\mathbf{u}\times\mathbf{B}) + \nabla\times(\frac{\eta}{\mu_0}\nabla\times\mathbf{B}) = \mathbf{0}}$$

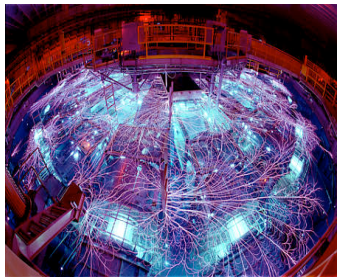
2D ↓

$$\mathbf{B} = \nabla\times\mathbf{A}$$

$$R_{A_z} = \frac{\partial A_z}{\partial t} + \mathbf{u}\cdot\nabla A_z - \frac{\eta}{\mu_0}\nabla^2 A_z + E_z^0 = 0$$



ITER



SNL Z-machine

Semiconductor Drift-Diffusion Model

(G. Hennigan, R. Hoekstra, J. Castro, D. Fixel, R. Pawlowski, E. Phipps, L. Musson, T. Smith, J. Shadid, P. Lin)

Electric potential

$$-\nabla \cdot \epsilon \nabla \psi = q\left(p - n + C\right)$$

$$\nabla \cdot \mathbf{J}_n - qR = q\frac{\partial n}{\partial t}$$

$$-\nabla \cdot \mathbf{J}_p - qR = q\frac{\partial p}{\partial t}$$

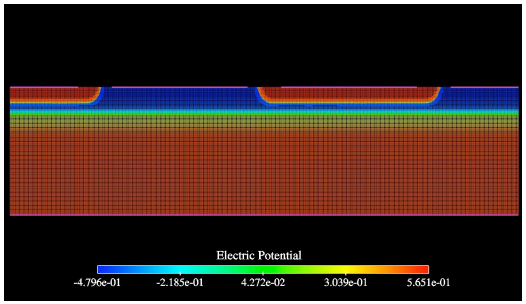
$$\mathbf{J}_n = -qn\mu_n \nabla \psi + qD_n \nabla n$$

$$\mathbf{J}_p = -qp\mu_p \nabla \psi - qD_p \nabla p$$

- ψ: electric potential
- n: electron concentration
- p: hole concentration
- C: doping profile
- R: generation-recombination term

Defect species: each additional species adds an additional transport-reaction equation

Stabilized FEM with Newton-Krylov solver with Trilinos Solvers (both MHD and drift-diffusion use same solver technology)

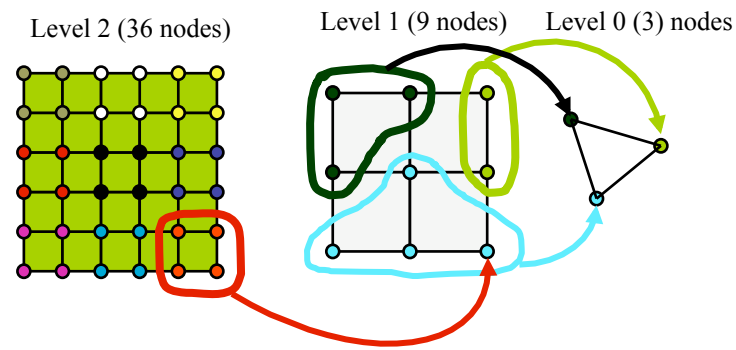


Trilinos ML Library: Algebraic Multigrid Preconditioners

(R. Tuminaro, J. Hu, C. Siefert, M. Sala, M. Gee, C. Tong)

- Aggressive coarsening with graph partitioner and pre-specified # of levels
  - Large difference in size between levels
  - Graph partitioner: serial for all levels, parallel for final level

- Aggregates to produce a coarser operator
  - Create graph where vertices are block nonzeros in matrix A_k
  - Edge between vertices i and j added if block $B_k(i,j)$ contains nonzeros
  - Decompose graph into aggregates
- Restriction/prolongation operator
- $A_{k-1} = R_k A_k P_k$



- Petrov-Galerkin smoothed aggregation for nonsymmetric matrices
  - Separate restriction smoothing
  - Local damping parameters

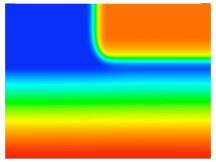
$$P_i = (I - \omega_i D^{-1} A)\hat{P}_i$$
$$R_i = \hat{P}_i^T (I - A D^{-1} \omega_i^{(r)})$$

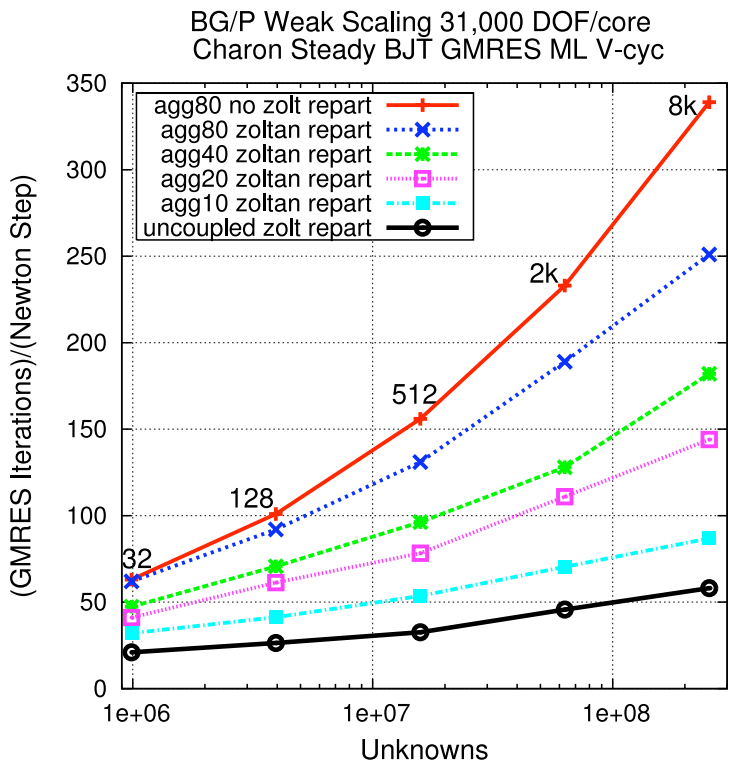
Coarsening Schemes

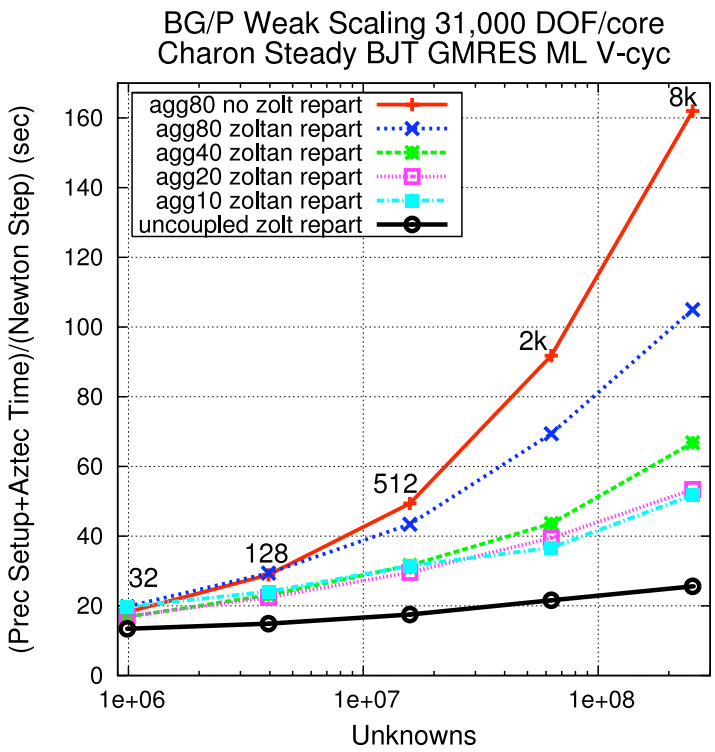
- Aggressive coarsening with graph partitioner
  - Serial graph partitioner (METIS) for each aggregation except for final which uses parallel graph partitioner (ParMETIS)
- (Aggressive) coarsening with graph partitioner and matrix repartitioning (Zoltan; RCB)
  - Graph partitioner (METIS) for all levels
  - Keep coarsening until matrix is below threshold size
  - More “mesh nodes” on cores; better quality aggregates
- Uncoupled aggregation with matrix repartitioning
  - Uncoupled: stencil is nearest neighbor, aggregates cannot span processes
  - Keep coarsening until matrix is below threshold size
  - Smaller difference in size between levels, e.g. ~9 for FEM 2D drift-diffusion
  - Better quality aggregates

BG/P Weak Scaling: Different Coarsening Schemes

- 2D BJT steady-state drift-diffusion
- 4-level aggressive coarsening V(1,1)
- METIS with matrix repartitioning, Uncoupled with matrix repart
- Scaled to 8192 cores and 250 million DOF; 31000 DOF/core



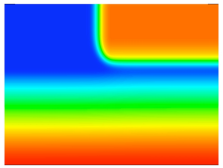




- Iteration Count
- GMRES
- CPU Time

BG/P Weak Scaling Study: 1-level vs. Multigrid

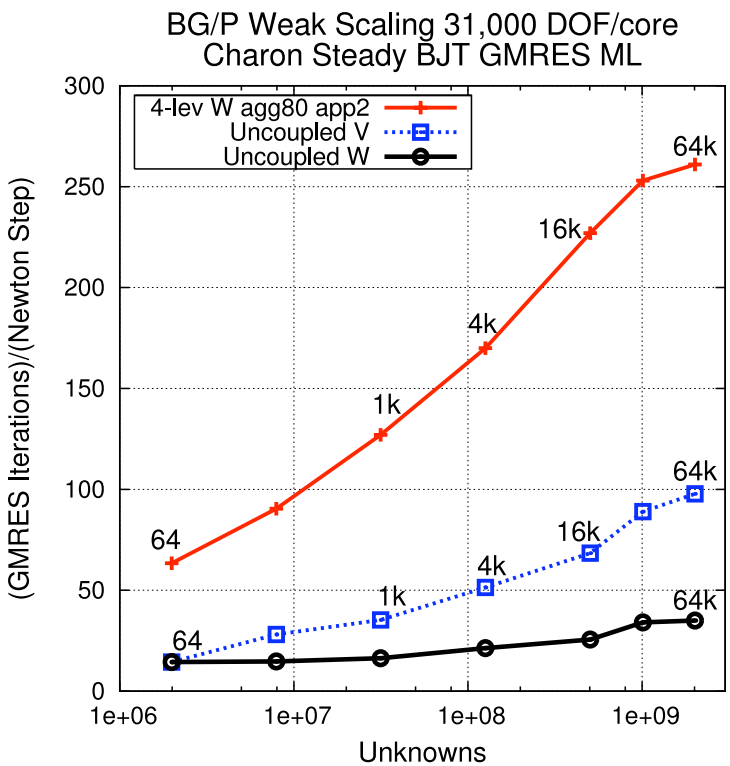
- 2D BJT steady-state drift-diffusion
- Uncoupled aggregation with matrix repartitioning
- Problem scaled to 8192 cores and 252 million DOF
- GMRES Krylov solver

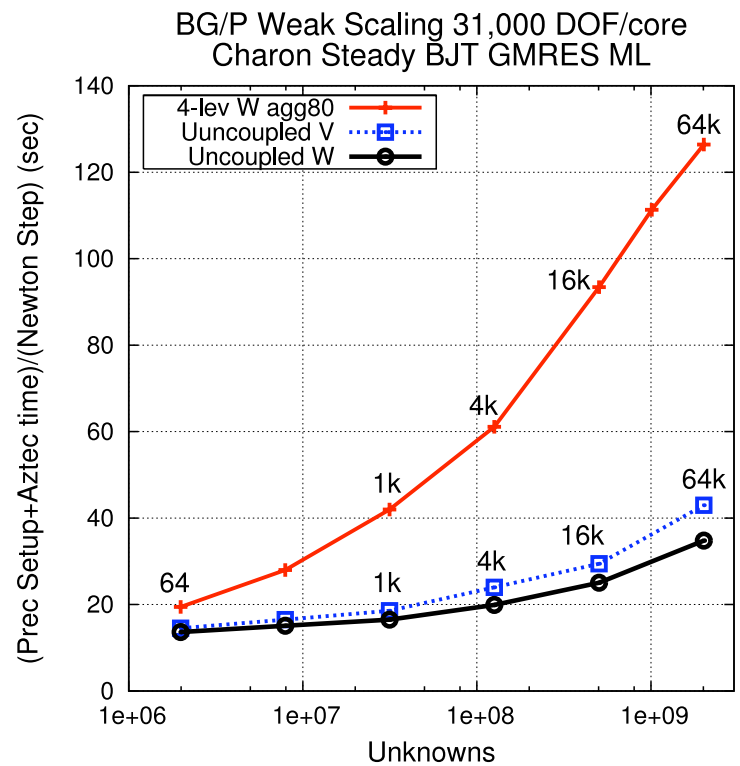


core	fine grid unknowns	1-level ILU		Uncoupled V-cyc		Uncoupled W-cyc	
		ave its per Newt step	time per Newt (s)	ave its per Newt step	time per Newt(s)	ave its per Newt step	time per Newt (s)
32	988533	214	55	21	13.5	14	13.1
128	3.95E+06	435	192	26	14.9	15	13.6
512	1.58E+07	859	697	33	17.5	16	15.6
2048	6.31E+07	1697	2634	46	21.6	20	18.1
8192	2.52E+08	3377	10559	58	25.6	25	22.6

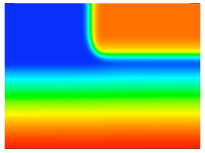
- Compared with 1-level preconditioner for 8192-core, 252 million DOF case
  - Uncoupled agg V-cyc reduces iterations by 182x, time by 412x
  - Uncoupled agg W-cyc reduces iterations by 422x, time by 467x

Reducing Iteration Count: Improved Aggregation

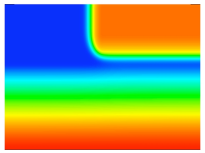


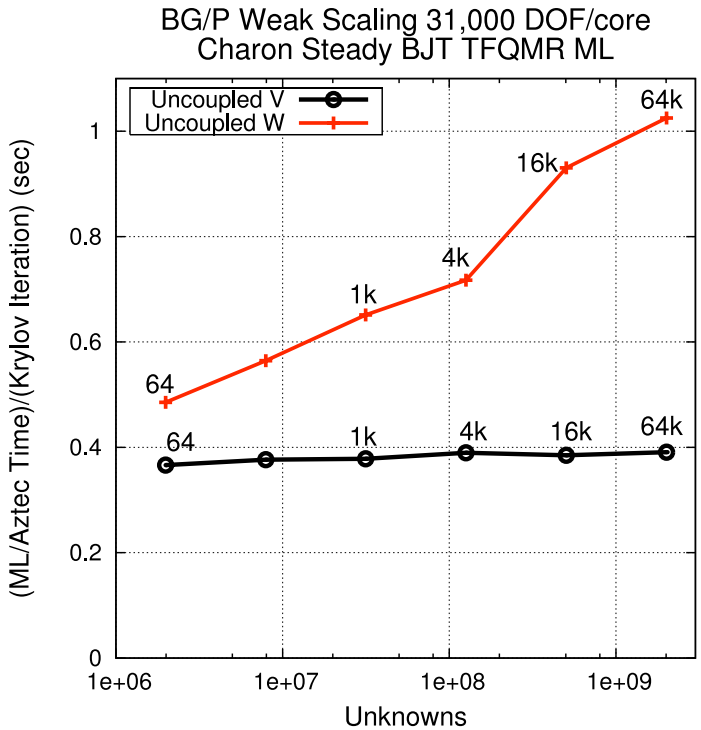


- Uncoupled with matrix repartitioning
  - more levels (up to 7)
  - better aggregates
- Significantly reduces iterations: W-cyc by ~8x, V-cyc by ~3x for 2 billion DOF for 64k
- Time reduction: W-cyc 3.6x 2 billion DOF, 64k

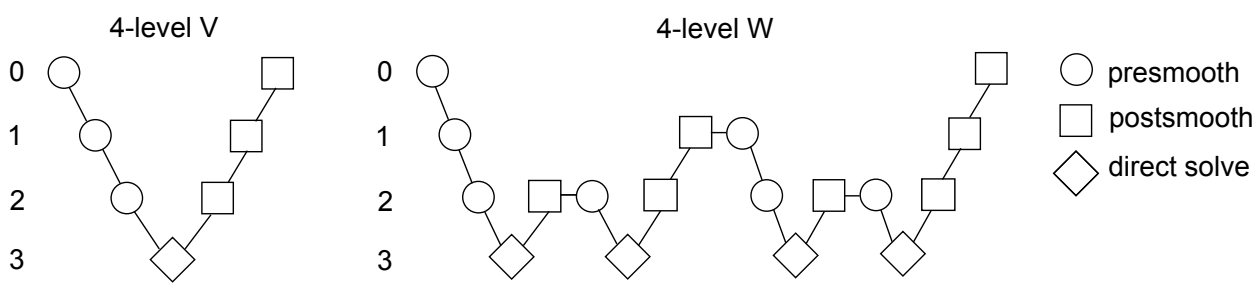


Uncoupled Aggregation: Time/iteration on BG/P



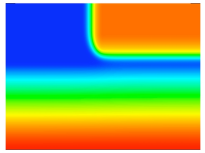


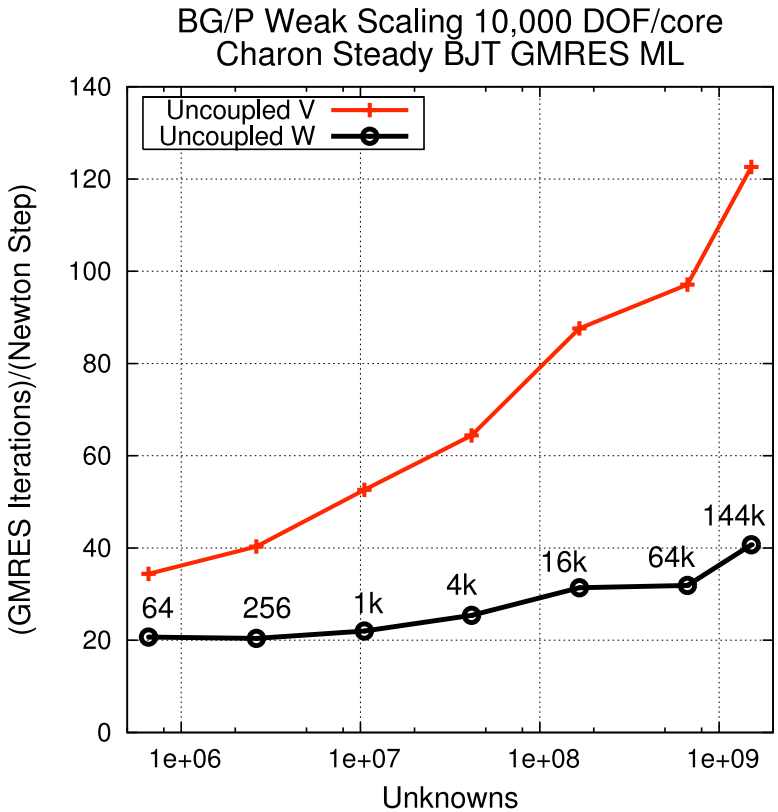
- TFQMR: time/iteration
- V-cyc time/iteration flat from 64 to 64k cores
- W-cyc time/iteration not doing well due to increase in work on coarse levels (7-lev)
- V-cyc and W-cyc about same total time; but W-cyc had fewer iterations/Newton

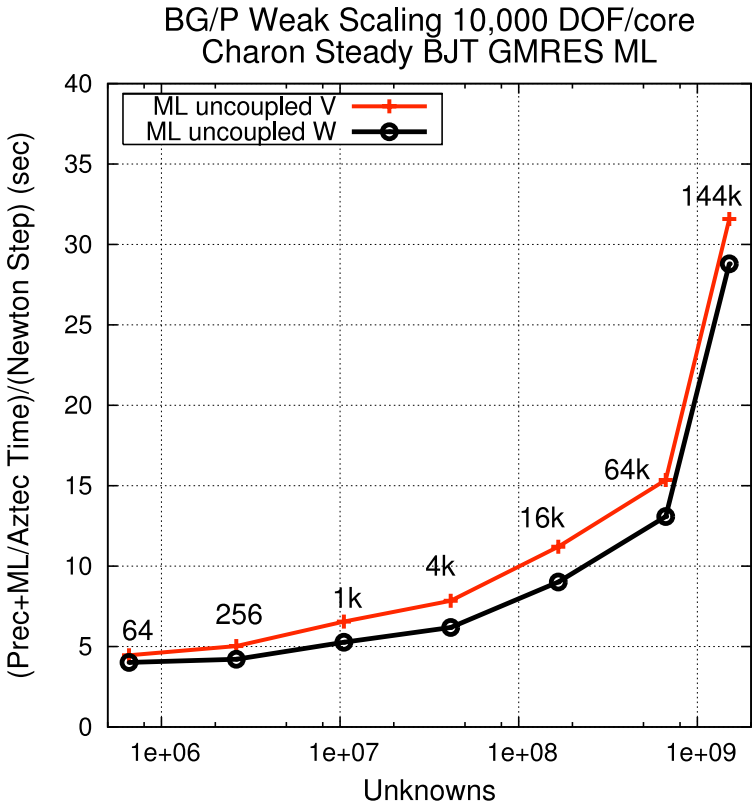


Weak Scaling to 147,000 Cores on IBM BG/P

2D Steady-state Drift-diffusion BJT

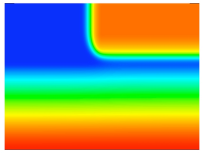


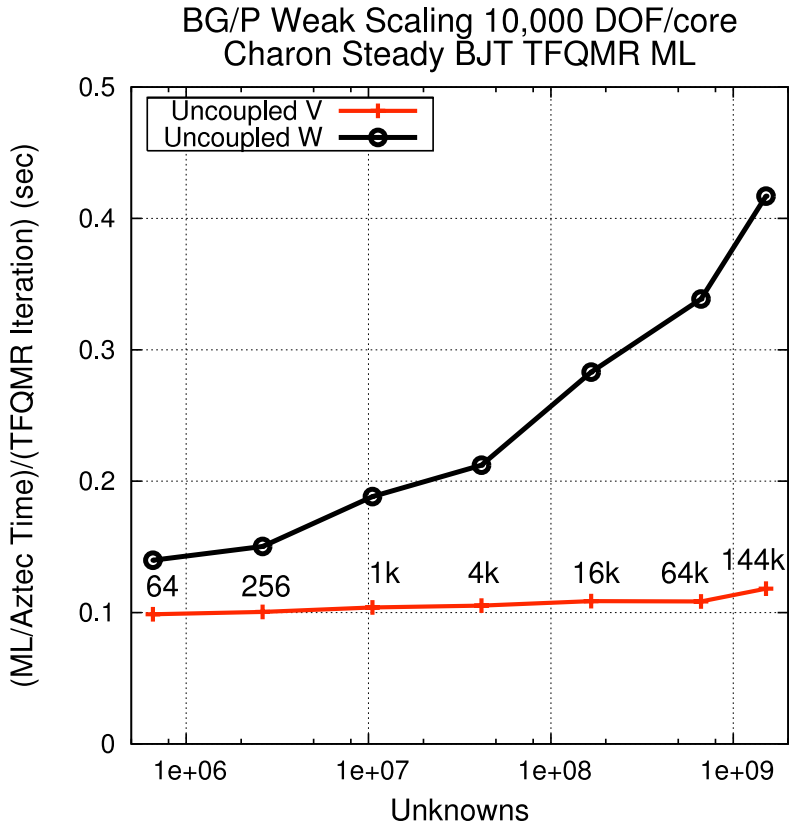




- 10,000 DOF/core; 1.47 billion DOF at 147,000 cores
- GMRES; Uncoupled aggregation with matrix repartitioning
- Problem size increased 2304x: W-cyc iter increased 2.0x; (prec setup+Aztec) time increased 7.2x

Weak Scaling to 147,000 Cores: Time/Iteration

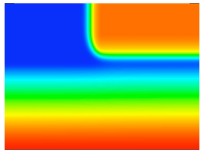




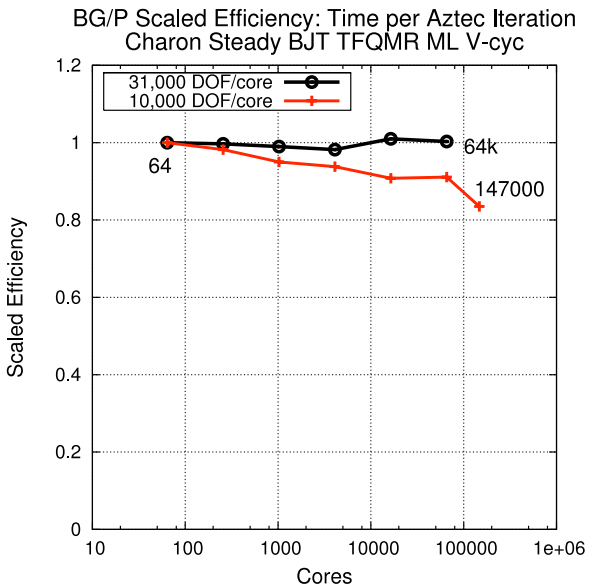
- 10,000 DOF/core; 1.47 billion DOF at 147,000 cores
- TFQMR uncoupled aggregation
- V-cyc: time/iteration
  - 64 to 64k cores increases 10% (1024x size increase)
  - 64k to 144k increases 9%
- W-cyc time rapidly increases due to larger amount of work at coarser levels

The authors thank the LLNL BG/P team for its help

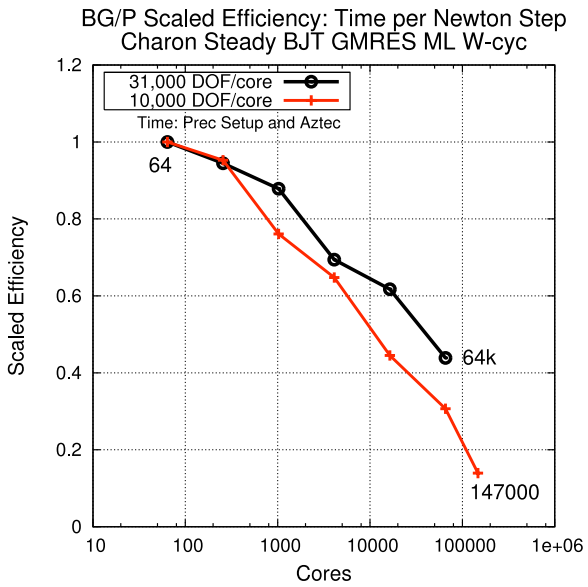
Overall Performance: Still Have More Work To Do



- Time per iteration scales well for V-cycle
- Next challenge to improve overall performance!
  - Improve preconditioner setup time
  - Improved repartitioning to minimize data movement in traversing mesh hierarchy and application of preconditioner
  - Eliminate re-computation of symbolic graph algorithms for projection and for matrix graphs (static meshes)
  - Work to obtain true h-independent iteration counts

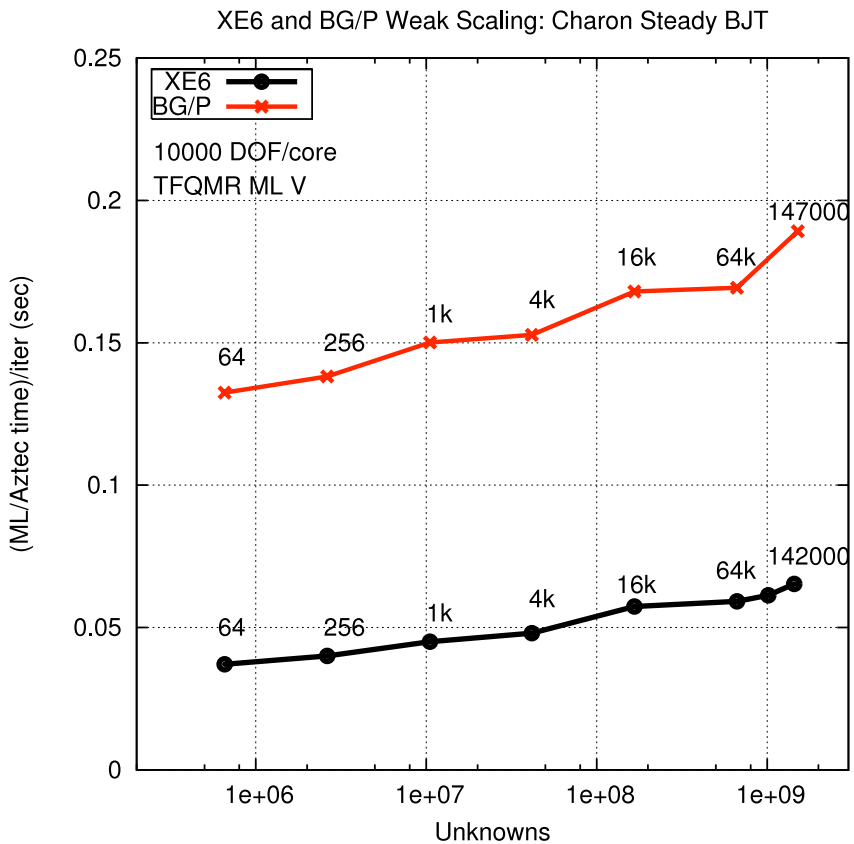


Time per Krylov Iteration



Linear Solve Time per Newton Step

Cray XE6 and IBM BG/P Weak Scaling

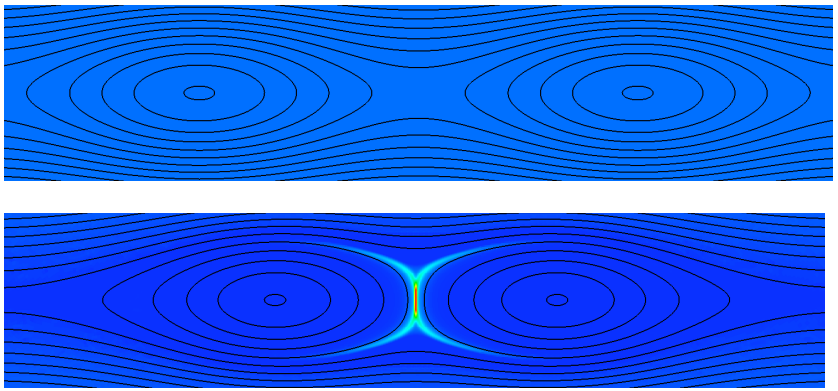


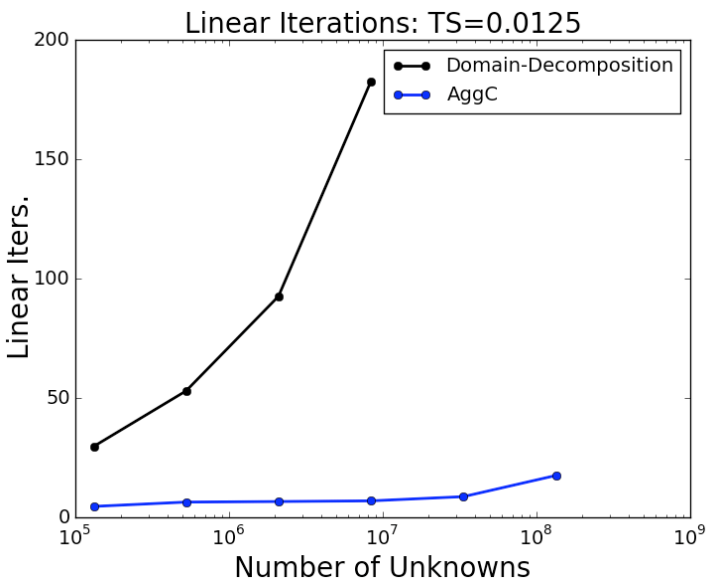
- Steady-state drift-diffusion BJT
- TFQMR time per iteration
- Cray XE6 2.4GHz 8-core Magny-Cours
- IBM Blue Gene/P 850 MHz quadcore PowerPC
- 10,000 DOF/core

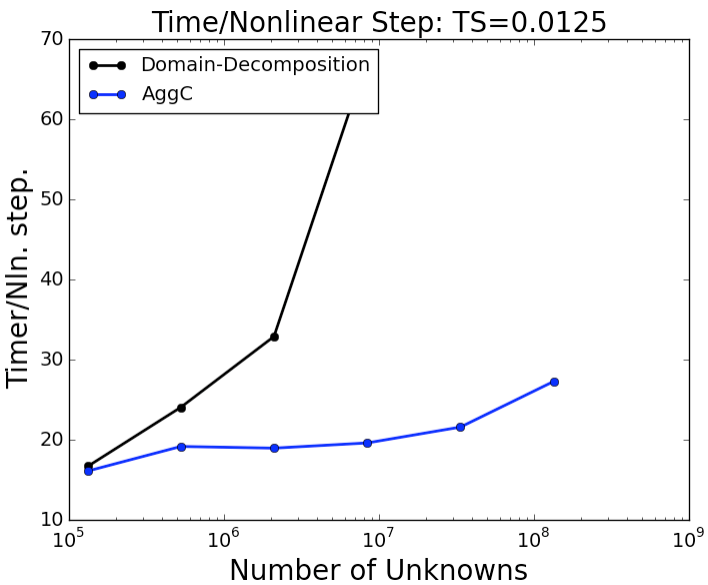
2D Resistive MHD: Island Coalescence Driven Magnetic Reconnection

(with R. Pawlowski, E. Cyr, L. Chacon)

- Half domain symmetry on [0,1]x[-1,1] with S=1e+4
- Linux cluster: Nehalem processors and InfiniBand interconnect

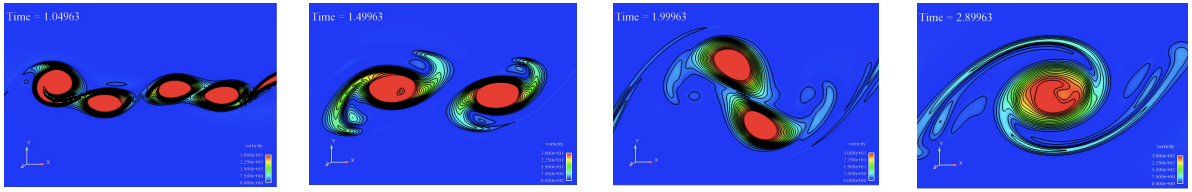


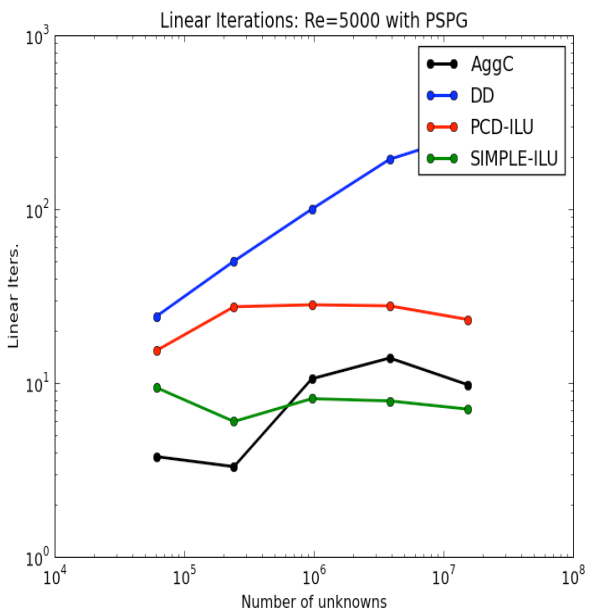


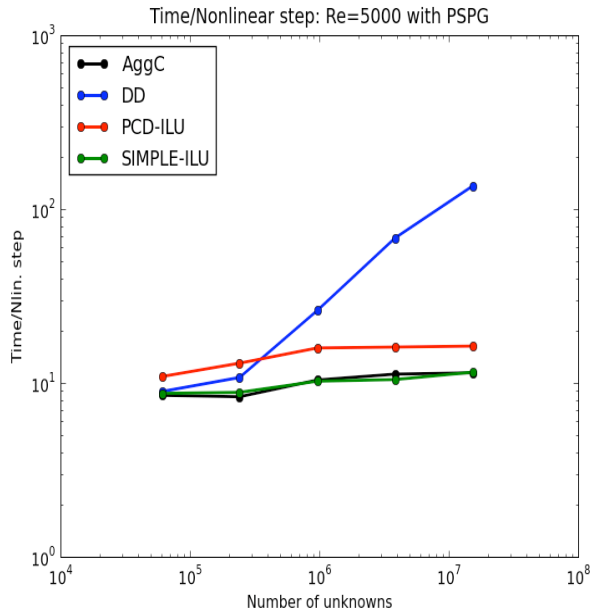


Transient Kelvin-Helmholtz CFD

(with Eric Cyr, R. Pawlowski)







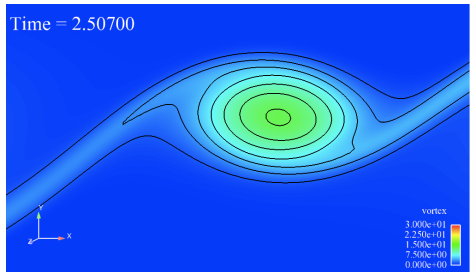
Kelvin-Helmholtz: Re=5000, Weak scaling at CFL=2.5

- Run on 1 to 256 cores – Nehalem cluster with InfiniBand
- Pressure - PSPG, Velocity - SUPG (residual and Jacobian)

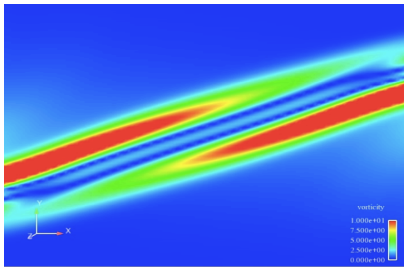
Transient Hydromagnetic Kelvin-Helmholtz Problem

(with R. Pawlowski, E. Cyr, L. Chacon)

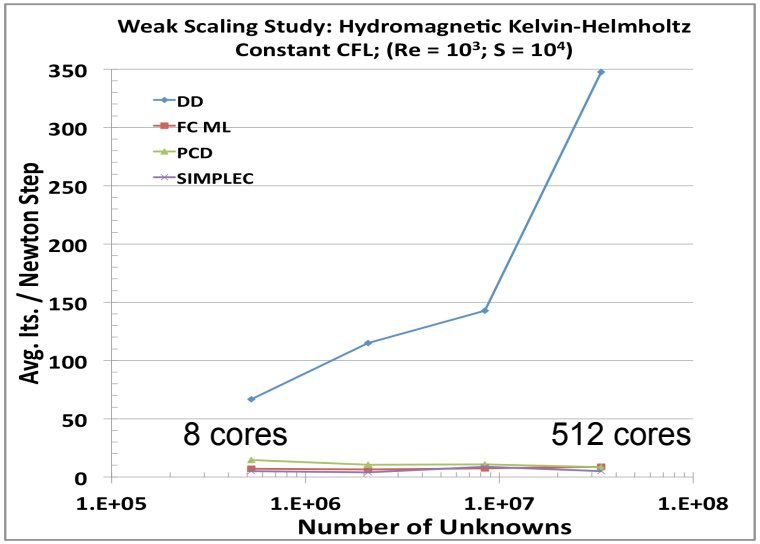
Re = 1000, S = 10,000, CFL = ~2.5

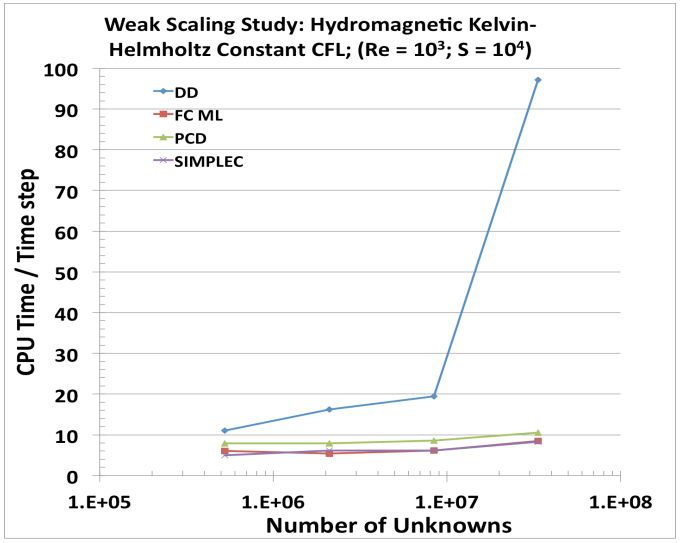


No magnetic field



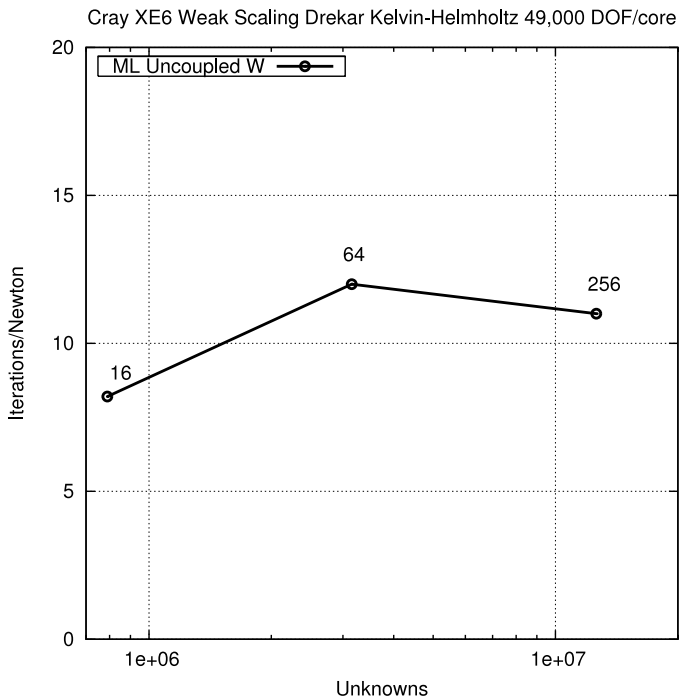
With magnetic field

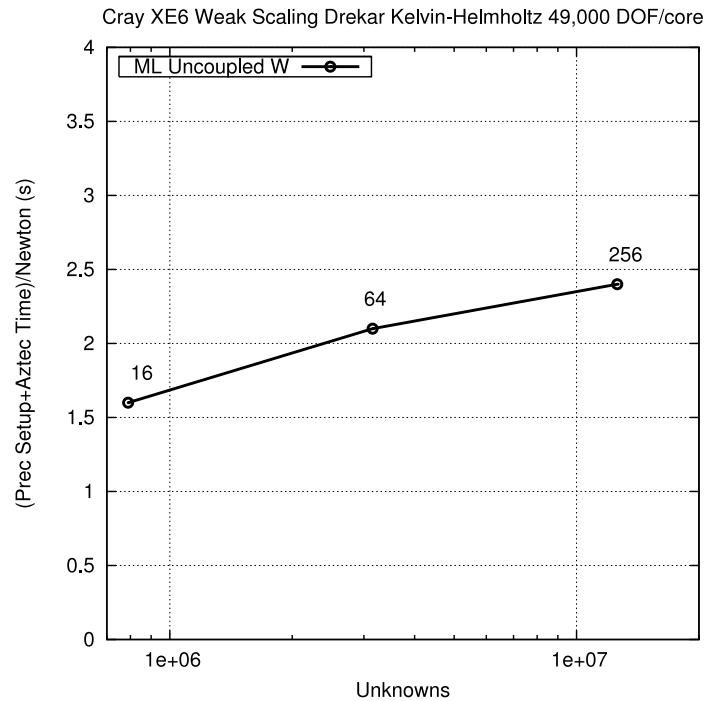




Linux cluster: Quadcore Nehalems with InfiniBand

Cray XE6 Preliminary Weak Scaling Kelvin-Helmholtz





- Re=5000, CFL ~2.5
- ML: uncoupled
- Cray XE6: dual-socket 2.4GHz 8-core Magny-Cours
- Next: Higher Re, MHD problems, larger problems

Concluding Remarks and Future Work

- Newton-Krylov/AMG methods can provide an effective, robust and efficient solution technology for a range of highly-nonlinear multiple-time-scale multiphysics systems.
- Massively parallel simulations on up to 140,000 cores
  - Time per iteration for V-cycle scales well
  - Next challenge to improve preconditioner setup and iteration count
- Future work: MHD problems at scale
- Issues
  - Strong convection effects, hyperbolic systems
  - Highly non-uniform FE aspect ratios
  - Hybrid parallelism; GPUs/accelerators

References:

[1] P. Lin, J. Shadid, M. Sala, R. Tuminaro, G. Hennigan, R. Hoekstra, “Performance of a Parallel Algebraic Multilevel Preconditioner for Stabilized FE Semiconductor Device Modeling,” Journal of Computational Physics, Vol 228, Issue 17, pp. 6250-6267, 2009

[2] P. Lin, J. Shadid, “Towards Large-Scale Multi-Socket, Multicore Parallel Simulations: Performance of an MPI-only Semiconductor Device Simulator,” Journal of Computational Physics, Vol 229, Issue 19, pp. 6804-6818, 2010

[3] J. Shadid, R. Pawlowski, J. Banks, L. Chacon, P. Lin, R. Tuminaro, “Towards a Scalable Fully-Implicit Fully-coupled Resistive MHD Formulation with Stabilized FE Methods,” Journal of Computational Physics, Vol 229, Issue 20, pp. 7649-7671, 2010

[4] P. Lin, “Improving multigrid performance for unstructured mesh drift-diffusion simulations on 147,000 cores,” submitted to the International Journal for Numerical Methods in Engineering, 2011

Acknowledgment:

The authors gratefully acknowledge funding from the DOE Office of Advanced Scientific Computing Research (ASCR) Applied Mathematics Program and DOE NNSA Advanced Simulation & Computing (ASC) program