Parallelization of DEBS

Sherwood Conference 2016

B.S. Cornille and C.R. Sovinec

University of Wisconsin-Madison

There is motivation for parallelizing DEBS

- Would like to reduce turn-around time for simulations.
  - Dr. Josh Reusch ran one simulation for about a year of real time for his thesis to produce an ensemble of sawtooth events.
- Would like to run simulations with greater resolution.

DEBS solves an MHD initial value problem

MHD Equations:

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{v} \times \mathbf{B} - \eta \mathbf{J}$$

$$\rho \frac{\partial \mathbf{v}}{\partial t} = -\rho \mathbf{v} \cdot \nabla \mathbf{v} + \mathbf{J} \times \mathbf{B} - \nabla P + \nu \nabla^2 \mathbf{v}$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{v}$$

$$\frac{\partial P}{\partial t} = -\gamma P \nabla \cdot \mathbf{v} - \mathbf{v} \cdot \nabla P + (\gamma - 1)[\nabla \cdot \underline{\underline{\mathbf{K}}} \cdot \nabla \left(\frac{P}{\rho} \right) + Q]$$

$$\mathbf{B} = \nabla \times \mathbf{A} \qquad \mathbf{J} = \frac{1}{\mu_0} \nabla \times \mathbf{B}$$

Assumed periodic cylinder geometry

- (r, θ, z) with $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$, and $0 \leq z \leq L = 2\pi R$

DEBS uses a spectral representation for θ and z

- DEBS geometry is periodic in θ and z, so a Fourier spectral method is used in these coordinates.
- Spectral methods have *geometric convergence* for *smooth functions*.
- Complex finite Fourier coefficients defined

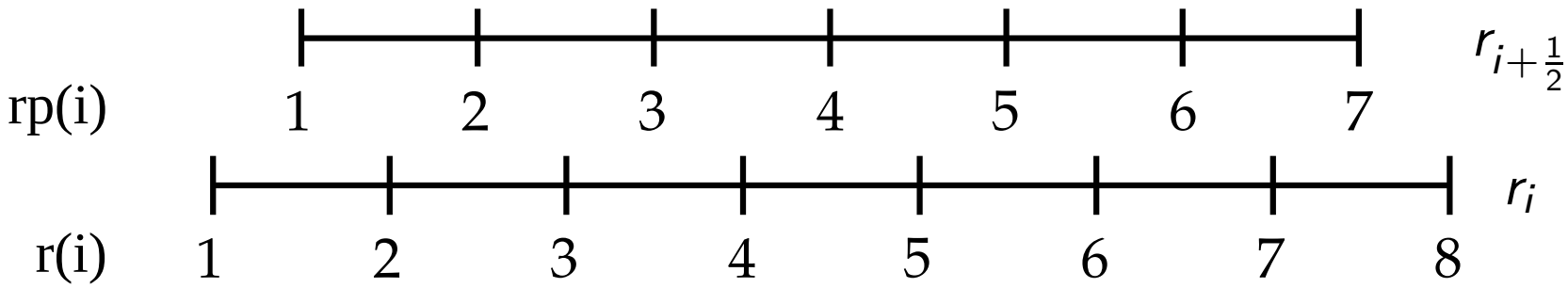
$$f_{m,n}(r) = \frac{1}{MN} \sum_{j=1}^{M} \sum_{k=1}^{N} f(r, \theta_j, z_k) e^{i(m\theta_j + n\frac{z_k}{R})}$$

- In a Fourier representation differential operations become arithmetic.

$$\left(\frac{\partial f(r,\theta,z)}{\partial \theta} \right)_{m,n} = -imf_{m,n}(r) \quad \left(\frac{\partial f(r,\theta,z)}{\partial z} \right)_{m,n} = -i\frac{n}{R} f_{m,n}(r)$$

DEBS utilizes a staggered mesh for finite differences in r

- DEBS uses a staggered mesh in r



- A_θ, A_z, B_r, v_r, J_θ, J_z, ρ, and P defined on r_i
- A_r, B_θ, B_z, v_θ, v_z, and J_r defined on $r_{i+\frac{1}{2}}$

The **curl** operator motivates the use of a staggered mesh

The evaluation of example terms of the **curl** operator on the DEBS mesh:

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

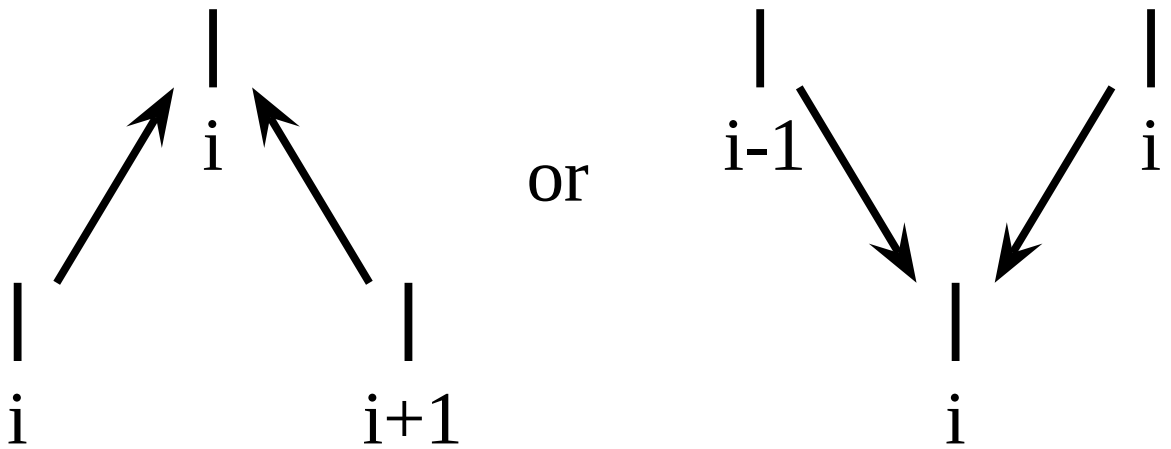
$$B_\theta = \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \qquad J_\theta = \frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r}$$

$$\left.\frac{\partial A_z}{\partial r}\right|_{i+\frac{1}{2}} \approx \frac{A_{z,i+1} - A_{z,i}}{\Delta r_{i+\frac{1}{2}}} \qquad \left.\frac{\partial B_z}{\partial r}\right|_{i} \approx \frac{B_{z,i+\frac{1}{2}} - B_{z,i-\frac{1}{2}}}{\Delta r_i}$$

$$B_{\theta,i+\frac{1}{2}} \approx -i\frac{n}{R}A_{r,i+\frac{1}{2}} - \frac{A_{z,i+1} - A_{z,i}}{\Delta r_{i+\frac{1}{2}}} \qquad J_{\theta,i} \approx -i\frac{n}{R}B_{r,i} - \frac{B_{z,i+\frac{1}{2}} - B_{z,i-\frac{1}{2}}}{\Delta r_i}$$

Data dependencies motivate the domain decomposition

Radial derivatives from **curl** evaluation can be pictorially represented:



DEBS uses a semi-implicit time advance

- Explicit methods are subject to a Courant-Friedrichs-Lewy (CFL) condition

$$\Delta t < \frac{\Delta x}{v}$$

- Fast MHD waves lead to too restrictive of time steps.
  - e.g. electrostatic, fast compressional, and shear Alfvén
- Implicit methods produce unconditional numerical stability for waves
  - No CFL restriction to time step
  - Expensive for nonlinear terms
- Like implicit methods, semi-implicit methods are numerically stable for very large time steps. However, semi-implicit operators can be designed to not couple Fourier components.

How a semi-implicit time advance can be developed

- A semi-implicit scheme is developed by subtracting a linear term that is designed to mimic a the nonlinear term in the equation from each side of a time advance
  - This term is treated once implicitly and once explicitly, e.g. [1]

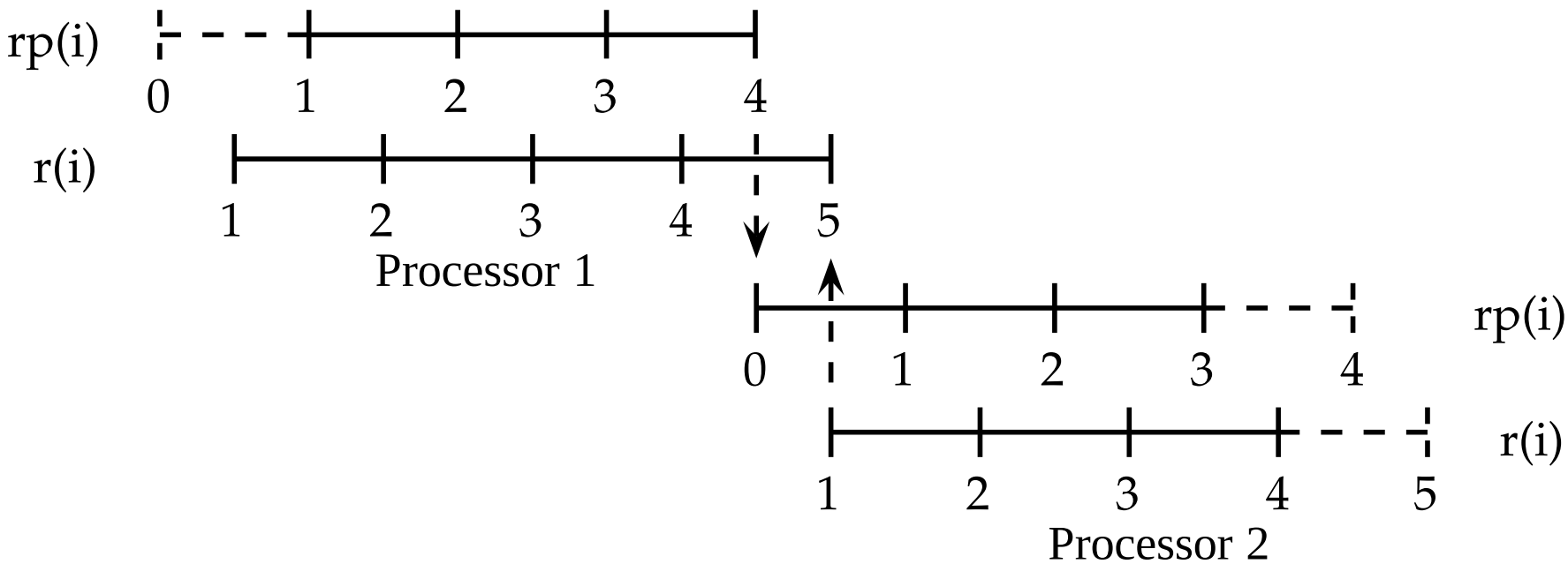
$$\mathbf{A}^{n+1}+\Delta t\eta_0\nabla\times\nabla\times\mathbf{A}^{n+1} = \mathbf{A}^{**}-\Delta t\eta\nabla\times\nabla\times\mathbf{A}^{**}+\Delta t\eta_0\nabla\times\nabla\times\mathbf{A}^{**}$$

- Each semi-implicit time advance in DEBS only requires inversion of a block-tridiagonal matrix for each Fourier component.
- In DEBS, waves are stabilized by the semi-implicit advance, but there is still a much less restrictive CFL condition from advection.
  - i.e. $\mathbf{v}\cdot\nabla\mathbf{v}$

[1] D. S. Harned and D. D. Schnack, "Semiimplicit method for long-time scale magnetohydrodynamic computations in 3 dimensions", Journal of Computational Physics **65**, 57 (1986)

Using the existing data structures, a radial mesh decomposition was developed

Using the dependencies for derivative calculation, a general communication scheme was designed:



How parallel performance is measured

Strong Scaling

- Constant problem size; increased computer resources
- Measured by speedup $S = \frac{t_{seq}}{t_P}$ or efficiency $E = \frac{S}{P}$
- Ideal case: efficiency of one for all P

Weak Scaling

- Problem size scales with increases computer resources
- Ideal case: constant execution time

Some definitions for performance testing of Parallel DEBS

Three inputs that set mesh size; nr, mt, and mz

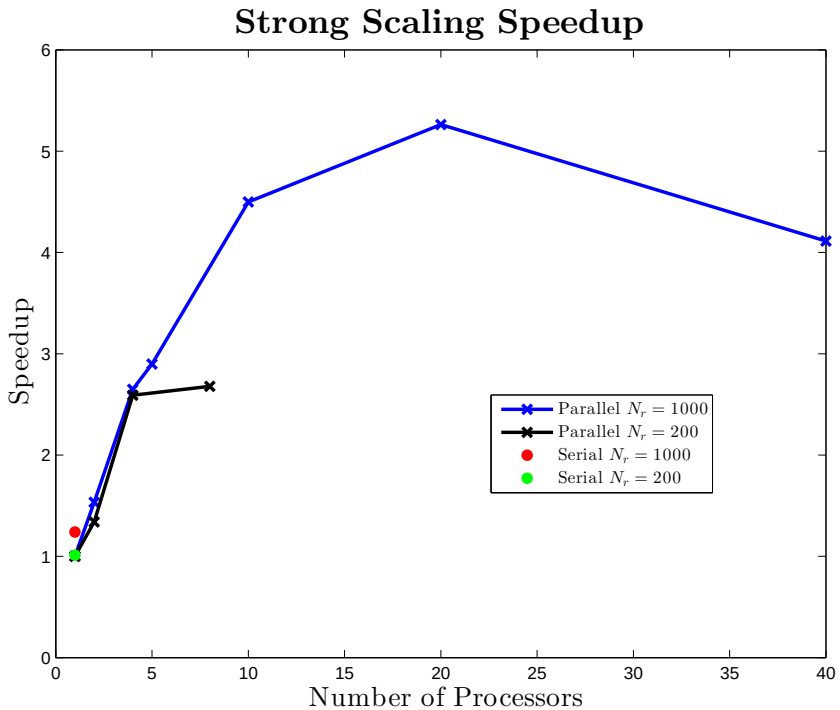
- nr is the array size in the radial direction for each processor
  - Global mesh size defined by nr_tot$=$(nr-1)$*$nprocs
    - nr_tot also denoted N_r
    - nprocs also denoted P
- $N_\theta = 2^{\text{mt}}$
- $N_z = 2^{\text{mz}}$

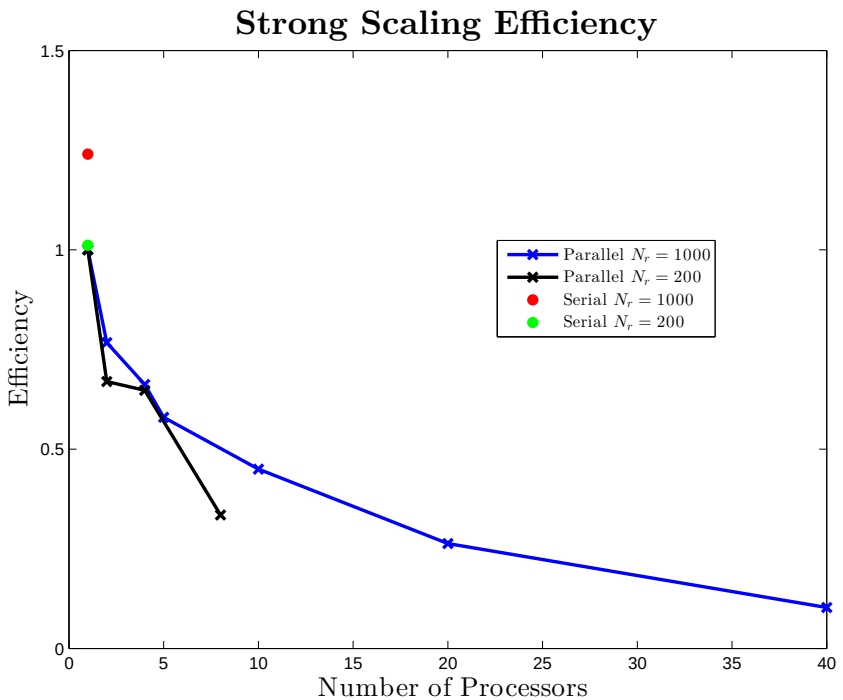
The high-performance computing (HPC) cluster of the UW-Madison Center for High Throughput Computing (CHTC) was used

- Two Intel Xeon E5-2670v2 processors per node; total of 20 processor cores
- 128 GB RAM per node
- Cluster connected by 56 Gbit/s Infiniband network
- OpenMPI version 1.6.4 used to compile DEBS, ScaLAPACK, and HDF5
- ATLAS version of LAPACK and BLAS used

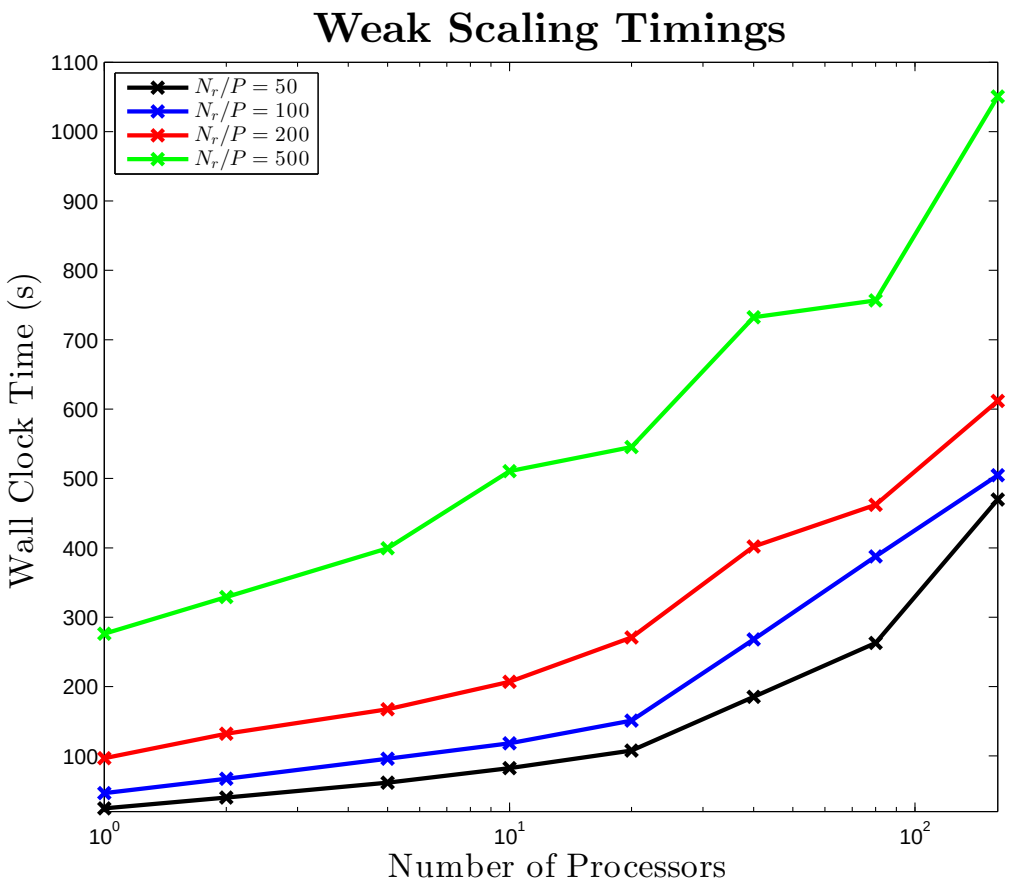
Strong scaling results

Two processors show speedup over serial DEBS, e.g. 17 min vs. 23 min, for equivalent problem sizes. Gains seen until about $\frac{N_r}{P} < 50$.





Weak scaling results



All sequences used mt $= 3$ and mz $= 5$ and were advanced 1000 time steps.

A relatively simple relation demonstrates how latency and bandwidth affect parallel banded-matrix solves

Total operation count of the ScaLAPACK routine used[2]

$$\varphi \approx 2k_l(4k_u+1)\frac{n}{p} + (4k_l+4k_u+1)r\frac{n}{p} + \left(\frac{32}{3}k^3 + 9k^2r + 6t_s + 4k(k+r)t_w\right)\lfloor \log_2(p-1)\rfloor$$

- p is the number of processors, n is the matrix size, r is the number of right hand sides, k_l is the number of lower matrix bands, k_u is the number of upper bands, and $k = \max(k_l, k_u)$
- t_s is the start-up time, in terms of floating-point operations (flops), of a single communication call, and t_w is the time, in flops, to transfer a word of data

[2] P. Arbenz et al., "A comparison of parallel solvers for diagonally dominant and general narrow-banded linear systems", Eidgenossische Technische Hochschule, Department of Computer Science, Institute of Scientific Computing (1998)

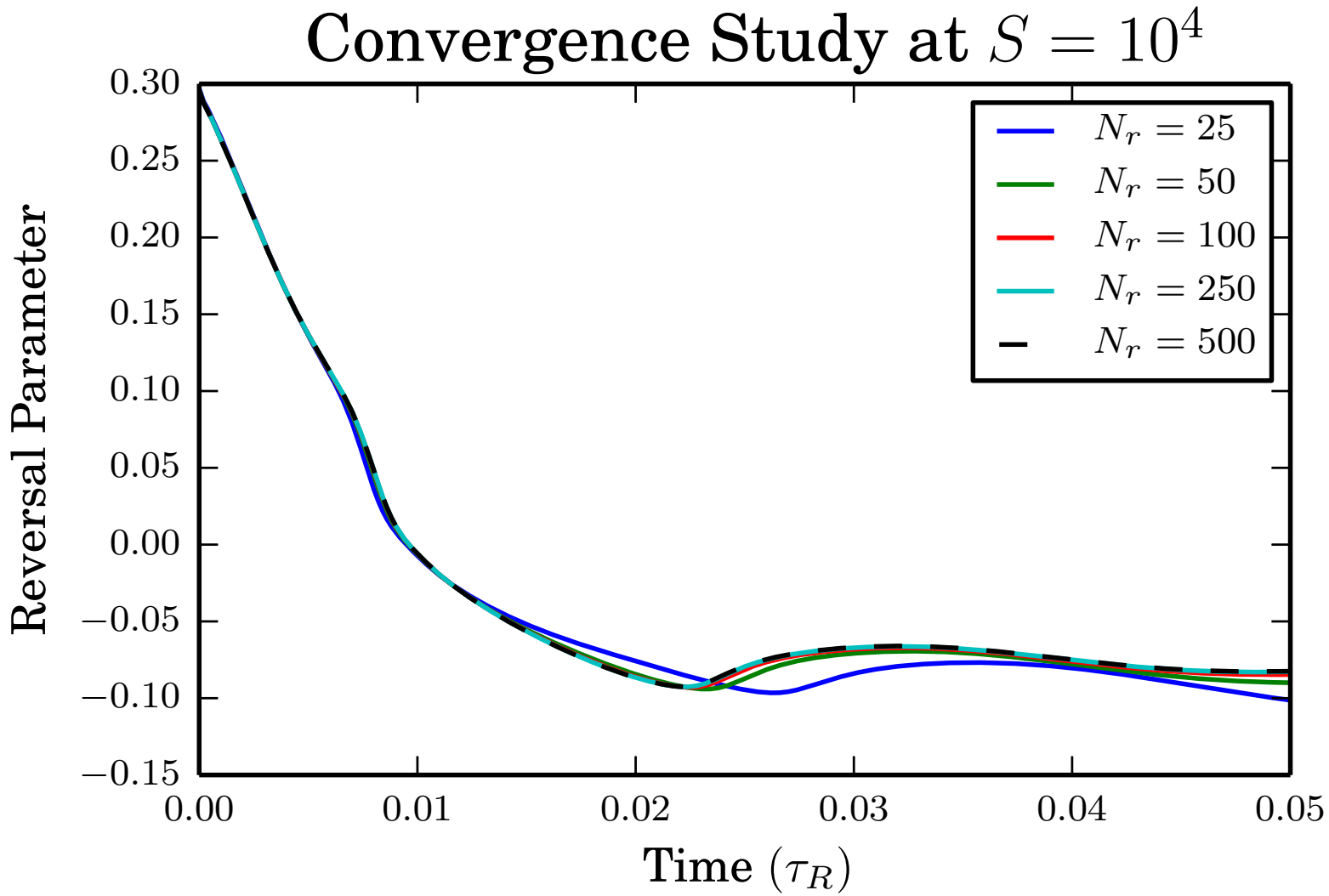
Parallel DEBS appears to be limited by latency

- Reasonable estimates for t_s and t_w are $t_s \approx 1000$ and $t_w \approx 10$
- Three matrix sizes solved in DEBS with $k_l = k_u = k = 1, 3, 5$ and $n = N_r, 2N_r, 3N_r$
- Using $k_l = k_u = k = 1$, and $n = N_r$ emphasizes the communication terms the most, giving a simplified operation count

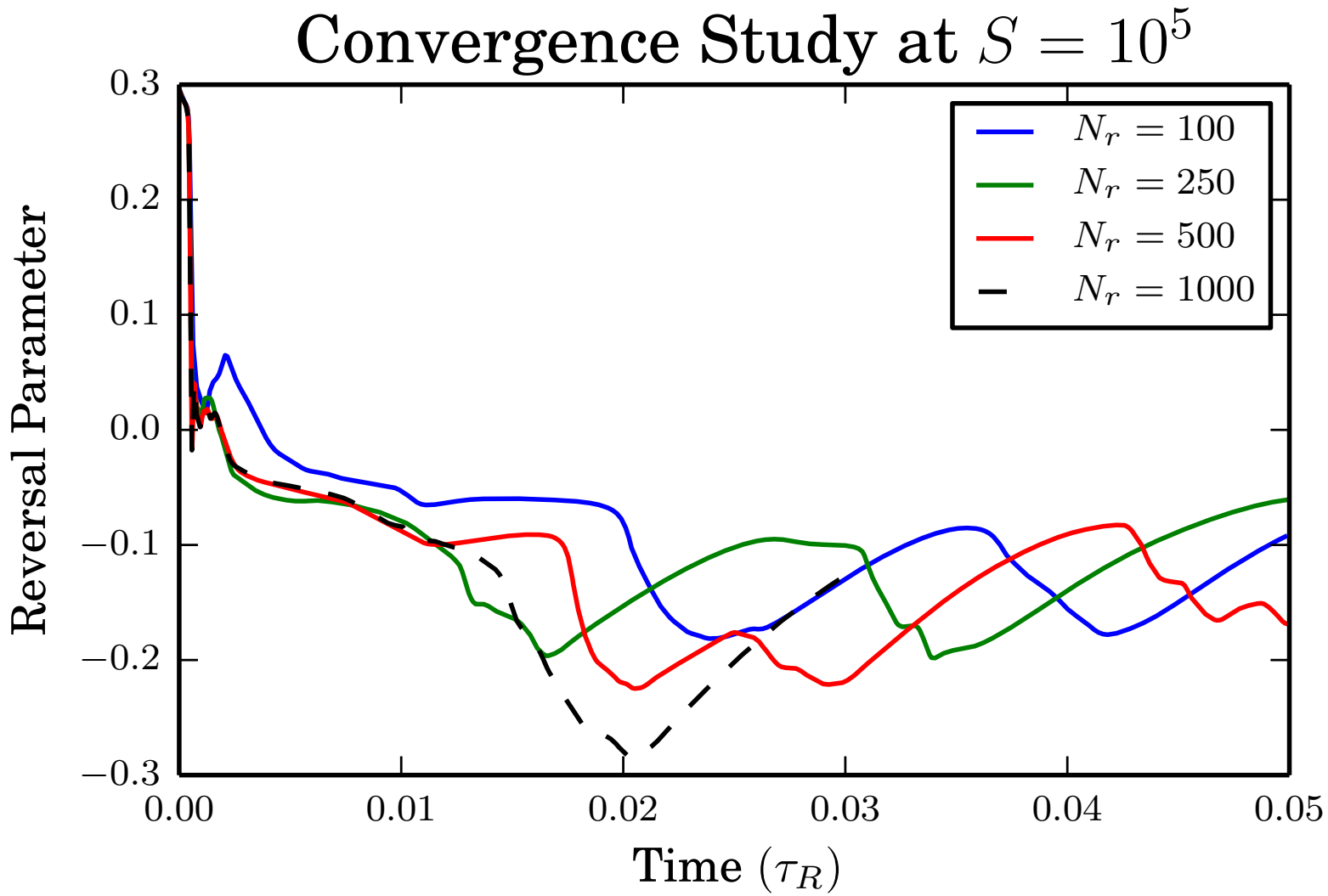
$$\varphi \approx 19\frac{N_r}{p} + (20 + 6t_s + 8t_w)\lfloor \log_2(p-1) \rfloor$$

- This shows communication dominated operation count, specifically t_s, for $\frac{N_r}{P}$ values studied
- t_s is different within a shared-memory node vs. over a network
- Parallel DEBS calculations should mostly be on single pieces of hardware

At modest Lundquist number numerical convergence requires modest radial resolution



At higher Lundquist number numerical convergence requires substantial radial resolution



Findings

- Gains have been made in turnaround time and practically available radial mesh sizes for DEBS simulations.
  - Parallel strong scaling is limited to $\frac{N_r}{P} \gtrapprox 50$.
  - Parallel execution should mostly be done on share-memory hardware.
- The improved radial resolution made accessible by Parallel DEBS will be especially important for high-Lundquist number simulations.
- The CFL condition for advection begins to become limiting at very high radial resolution.