A new planet in 7 days: A job for BlueGene/L

A.Wray, J.Jimenez, P.Moin, L.Shunn,
A.Mani, S.Jones, F.Ham, G.Iaccarino
29 August 2006

Center for Integrated Turbulence Simulations
Stanford University
A new planet in 7 days: A job for BlueGene/L

i. INTRODUCTION
The question is whether turbulence in a protoplanetary disk, either initially created by magnetic instabilities or by the accretion itself, would self sustain by drawing energy from the shear.

Nearly all models of planetary formation assume turbulent mixing and transport within the disk.

The goal is to simulate a turbulent disk at a high Reynolds number and observe the corresponding energy growth ($Re_\lambda \sim 200$).
What is turbulence?

• **Turbulence**: a flow regime characterized by chaotic, stochastic motion and rapid variation of pressure and velocity in both space and time. Flow that is not turbulent is called laminar flow.

• The dimensionless **Reynolds number**, \( Re = \frac{uL}{\nu} \), characterizes whether the flow is laminar or turbulent.
Homogeneous isotropic turbulence is an idealization of real turbulent flow, under the assumption that the motion is governed by a statistical law that is invariant for arbitrary translation (homogeneity), rotation or reflection (isotropy) of the coordinate system.

Kolmogorov (1940s) first derived a statistical theory to predict the behavior of turbulent motion without external forcing (natural decay).
Isotropic turbulence

Kolmogorov $k^{-5/3}$ energy cascade

The energy in unforced isotropic turbulence decays with time
Rapid distortion theory

- Introduced by Batchelor (1950s) to study the evolution of turbulence under external forcing (shear, gravity, etc.)
- Based on a linearization of the governing equation of motion, thus strictly valid only for small perturbation/distortion (with respect to the flow time scales)
- The energy spectrum in RDT evolves according to a close-form equation
- Provides useful insights in the physics and a solid ground for developing theories and phenomenological models
Rapid distortion theory

- Analytical expressions for the energy spectra provide information about the decay as a function of the rotation and shear.
- Detailed analytical expressions for turbulent quantities derived by Kassinos et al. (2006) allow to infer the presence of turbulence structures (elongated streaks, columns, etc.)
Previous work

• Numerous **2D simulations** of this flow, and some **3D simulations**
• In all previous simulations, energy decayed
• **2D flows** lack most turbulence generation mechanisms
• Previous **3D simulations** (and experiments) have relatively **low** $\text{Re}_\lambda \sim 50$
• DNS of Kaneda *et al.* (2003) $\text{Re}_\lambda = 1000$, but doesn’t include **largest structures**
A new planet in 7 days: A job for BlueGene/L

ii. BLUEGENE / L
BlueGene/L

- DOE/IBM supercomputer at LLNL
- Installation Year: 2005
- #1 on Top 500 list: 280 TFLOPS
- Theoretical peak: 370 TFLOPS
- Compute nodes: 65536 (131072 CPUs)
- I/O nodes: 1024
- Memory: 32768 GB
BlueGene/L: architecture

System
- 64 cabinets
- 65,536 nodes
- (131,072 CPUs)
- (32 x 32 x 64)
- 180/360 TF/s
- 32 TiB*
- 1.5 MW
- 2,500 sq.ft.
- MTBF 6.16 Days

Cabinet
- 2 midplanes
- 1024 nodes
- (2,048 CPUs)
- (8 x 8 x 16)
- 2.9/5.7 TF/s
- 512 GiB* DDR
- 15-20 kW

Node Card
- 16 compute cards
- 0-2 I/O cards
- 32 nodes
- (64 CPUs)
- (4 x 4 x 2)
- 90/180 GF/s
- 16 GiB* DDR
- 500 W


Compute Card or I/O Card
- FRU (field replaceable unit)
- 25 mm x 32 mm
- 2 nodes (4 CPUs)
- (2 x 1 x 1)
- 2 x (2.8/5.6) GF/s
- 2 x 512 MiB* DDR
- 15 W

Compute Chip
- 2 processors
- 2.8/5.6 GF/s
- 4 MiB* eDRAM

User logs in here

(compare this with a 1988 Cray YMP/8 at 2.7 GF/s)
A new planet in 7 days: A job for BlueGene/L

iii. COMPUTATIONAL SETUP
Keplerian turbulence

- Protoplanetary disks need to be turbulent to form planets
- Simplest model is rotating, sheared box of turbulence
Governing equations

- continuity
  \[ u_{i,i} = 0 \]

- momentum
  \[ u_{i,t} + (u_i u_j)_{,j} + p_{,i} = \nu u_{i,jj} + 2\epsilon_{ij3}\Omega u_j \]
To boldly go ...

- Box size: $L_y = L_z = 2\pi$, $L_x = 8\pi$
- Keplerian turbulence: $S/\Omega = -3/2 = \text{(shear)/(rotation)}$
- $\text{Re}_\lambda$ (low=80, high=220) = (inertial force)/(viscous force)
- $\text{Ro}$ (high, med, low) = (inertial force)/(Coriolis force)
- $L_y/L = 75$
Spectral code

• based on the isotropic turbulence code written in the Vectoral language by Alan Wray and Robert Rogallo*

• modifications for shear and rotation using Rogallo’s transformation (see next slide)

• converted to C for compiling on BG/L

Rogallo’s transformation

- Computational box moves with the mean flow and undergoes shear to allow periodic boundaries in the $x_2$-direction (required for spectral discretization).
- Occasionally, when the “tilt” becomes excessive, the data must be remapped to an orthogonal grid.
Solution algorithm (1 of 2)

- Start in Fourier wave space
- Inverse FFT to get all required quantities in physical space (velocity and first derivatives)
- Compute non-linear terms, $(uu_x)$ etc. in physical space
- While in physical space, compute maximum velocity and limit time step based on CFL criteria
- FFT back to wave space and complete time derivative of velocity, including projection to eliminate dilatation
- Advance in time using low-storage 4th-order RK integration
Solution algorithm (2 of 2)

• For efficient FFT/iFFT, entire lines of data in the transform direction must be in local processor memory.

• To accomplish this, the data is managed in “pencils” that span the entire box in each direction (8192, 2048, 2048 in x, y, and z respectively).

• To get the data from the condition where it is penciled in y (initial state) to penciled in x or z, basically all data must be transferred between processors - total communication scales linearly with problem size.

• This is unlike solvers based on domain decomposition techniques where only surface data is transferred, and the communication scales sub-linearly.
A new planet in 7 days: A job for BlueGene/L

iv. COMPUTING AND DATA PROCESSING
Simulations

- Grid size: 8192x2048x2048 ~ 34 billion grid points
- Ran on 65536 CPUs (half of the machine)
- 10-20 Mhours/CPU
- FFT transforms require large data (full volume) communication multiple times per time-step

Data processing

- Each saved data field: 384 GB
- Inverse FFT on entire dataset must be performed to visualize solution in physical space
The machine remained stable for the entire week and 60% parallel efficiency was achieved (spectral code with multiple volume data communication in each timestep)
Post-processing

- Output files contain **Fourier coefficients** of velocity field

- **Inverse FFT** must be performed in order to visualize data in physical space

Each CPU writes 16x8x2048 datafile
1. Each processor reads in one \textit{y,z-plane} of data
2. 2-dimensional \textit{inverse FFT} is performed
3. Each processor writes a temporary file containing transformed plane of data
4. When all \textit{y,z-transforms} are complete, each processor reads one \textit{x-plane} of data
5. \textit{Inverse transform} is performed in \textit{x-direction}
6. Final data files are written
1. Each processor reads in one y,z-plane of data
2. 2-dimensional inverse FFT is performed
3. Each processor writes a temporary file containing transformed plane of data
4. When all y,z-transforms are complete, each processor reads one x-plane of data
5. Inverse transform is performed in x-direction
6. Final data files are written
1. Each processor reads in one $y,z$-plane of data
2. 2-dimensional inverse FFT is performed
3. Each processor writes a temporary file containing transformed plane of data
4. When all $y,z$-transforms are complete, each processor reads one $x$-plane of data
5. Inverse transform is performed in $x$-direction
6. Final data files are written
Visualization

Spectral filtering
• “Coarse” representation of simulation is produced by discarding high-wavenumber velocity components

Sub-domain decomposition
• Sub-volumes of arbitrary size and location can be independently designated and prepared for visualization
• Allows user to explore the data by examining a range of resolutions and locations
A new planet in 7 days: A job for BlueGene/L

v. RESULTS
Isotropic initial condition

\[ \text{Re}_\lambda = 220 \]
\[ L_y = L_z = 2\pi \]
\[ L_x = 8\pi \]
\[ \frac{L_y}{\mathcal{L}} = 75 \]

dataset can serve as I.C. for future LES studies
Range of scales
Keplerian turbulence
Energy spectrum

contours of $k_1 E(k_1)$

High-Ro (slow rotation)  Low-Ro (fast rotation)
not all scales are dying!

Low-Ro (fast rotation)
Energy spectrum

not all scales are dying!

Low-Ro (fast rotation)
Nonlinear structures: “streaks”

“I’m not dead yet.”
Nonlinear structures: “streaks”

Rossby number:
\[ \text{Ro} = \frac{u'}{S \mathcal{L}} \]

“Streaks” in wall flow:
\[ \text{Ro} \approx 10^{-1} - 10^{-3} \]
\[ \text{Re} > 100 \]

Astrophysics:
\[ \alpha = \frac{(u' \mathcal{L})}{(S \mathcal{L}^2)} = \text{Ro} \approx 2 \times 10^{-3} \]
Conclusions

- 7 days on BGL and still **no earth**
- Now we know **what** to look for and **where**
- Can probably use **smaller boxes** or with **LES**
Conclusions

• 7 days on BGL and still no earth
• Now we know what to look for and where
• Can probably use smaller boxes or with LES