The Challenges of Developing a High-Resolution Climate Model for the IBM:BGQ Platform

Chris Kerr and Shian-Jiann Lin
Geophysical Fluid Dynamics Laboratory

Bob Walkup
IBM Watson Research

V. Balaji
Princeton University

March 6-7, 2013

Geophysical Fluid Dynamics Laboratory

Experiments

- **Held-Suarez**: Simplified physics, dry-dynamics
- **Doubly Periodic Experiments**: Full-model physics, no-land, no-ice, prescribed SST's
- **Aquaplanet**: Full-model physics, no-land, no-ice, prescribed zonally symmetric SST's
- **AMIP** (Atmospheric Model Intercomparison Project): Full-model physics, land, prescribed SST's

Climate Studies

- **Purpose of study**:

  Provide fundamental understanding of role of clouds in dynamics of climate

- **Previous AMIP studies**:

  100km, 50km, 25km, and 12km resolutions

- **Proposed AMIP study**:

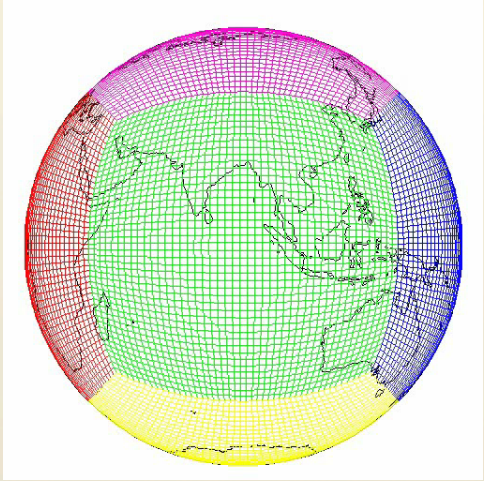
  3.5km resolution, 41 vertical Levels

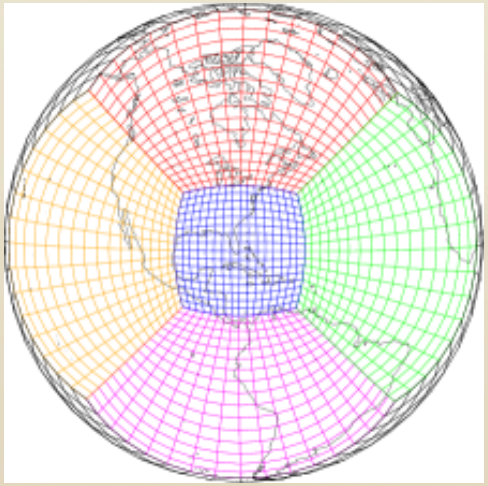
Components of Climate Model

- Infrastructure (GFDL)
- Atmosphere:
  - Dynamics (GFDL)
  - Physics (GFDL, others)
- Ocean (GFDL)
- Ice (GFDL, LANL)
- Chemistry (NCAR)

Cubed-Sphere Atmospheric Dynamics

- Fully compressible non-hydrostatic equations
- Mapping of cube to sphere
- Ideal [x,y] decomposition
- Finite-volume advection
- Grid nesting and stretching

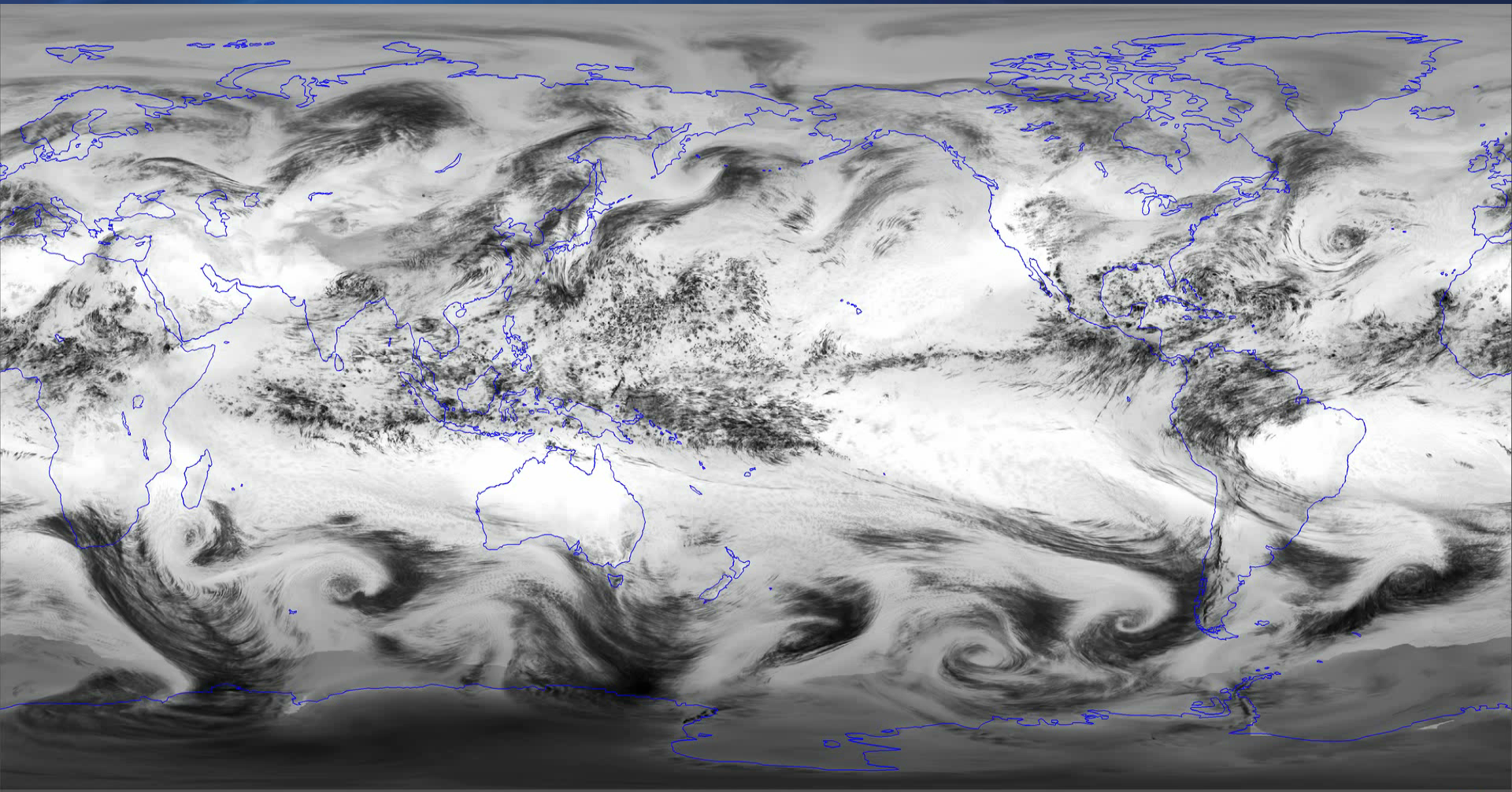




Status of Climate Experiments

- Experiments begin on May 1, 2005; run for several model months
- Tune model: **moisture** (precipitation), **heat** (OLR), …, tropical cyclones
- Interaction of dynamics and physics at 3.5km resolutions (significant changes in both)

OLR: September 2-30, 2005



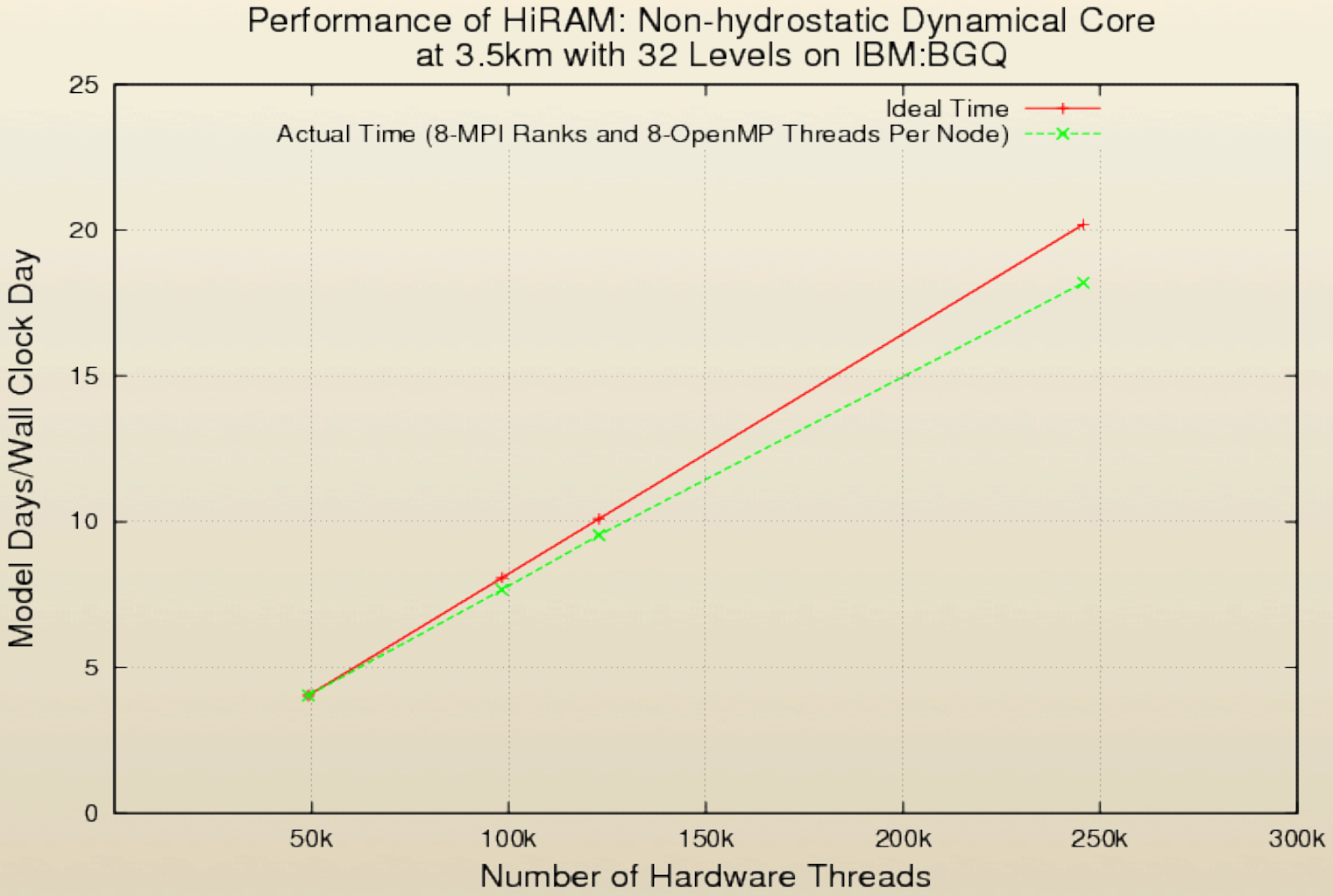
Computational Infrastructure

- Code written by GFDL scientists in FORTRAN 90 with MPI, and NetCDF
- 0.75M lines of executable code
- Software improvements over last 5-years: BGP and BGQ
  - OpenMP at high-level: atmosphere, land, …
  - Eliminated global domain arrays
  - Improve communication and I/O scaling

Computational Performance

- 3.5km executes on BGQ on 15,360 MPI ranks: (8-MPI Ranks/node and 8-OpenMP threads/rank)
- 9 model days per wall clock day
- Generates 250GB history files/model day
- History files downloaded and post-processed at GFDL
- 0.25M lines of post-processing scripts

Computational Performance



Computational-Science Challenges

- Model setup: (land-sea-ice masks, initial conditions, …)
- Issues with performing climate science at these scales:
  - Significant development time (2-day segments) months to equilibrate model
  - Dedicated resources needed for climate experiments
- Storage: 250GB/day

Computational Findings

- Improve performance on Mira to guide development
- Standalone versions of: atmosphere, physics, land, coupler, I/O, communications,… for performance testing
- Recurring theme,… code does not scale beyond, … 15,000 MPI-Ranks. 15% of time in I/O routines

Computational Findings

- HPM metrics, general characteristics:

  40% FPX-60% FXU. 10B/cycle (not bandwidth limited)

- Model timer counters: Atmosphere=75%, Physics=15%, land and coupler=10%. (70% compute and 30% communication)

- Use line and function-level timing data to direct optimizations. Identified:
  - Regions for OpenMP threading
  - Eliminate data copies at statement-level and across subroutine boundaries

Computational Challenges

- Future compute requirement: **3 model-years per wall clock day. Getting 9 model- hours per wall clock day**
- **Hundred-times** improvement to achieve required throughput
- Increase parallelism: blocking and task
- Storage/Post-processing: 2PB of data per scenario (20-years)

Summary

- Successful completion of Held-Suarez, Doubly-Periodic, and Aquaplanet Experiments
- AMIP experiments running at 3.5km; producing valuable scientific results
- Improve software performance on current system

Questions?

chris.kerr@noaa.gov