Recent Advancements in High Resolution Climate Modeling

14th Workshop on the use of HPC in Meteorology
1 November 2010

Dr. Richard Loft
loft@ucar.edu
Technology Development Division
Computational and Information Systems Laboratory
National Center for Atmospheric Research
Why High Resolution? Resolving Ocean Mesoscale Eddies

1° Ocean component of CCSM (Collins et al, 2006)

0.1° Eddy Resolving (Maltrud & McClean, 2005)
Ocean-Atmosphere Interactions: North Atlantic Winter Storm Track

- Stronger SST gradient
- Heavier precipitation

0.5° atm + 0.1° ocn
Engineering Petascale Software

In modeling we dream of speed and elegance... but if you're not careful... a different result
CESM: New Modeling Capabilities for Ultra-High Resolution Simulations

Including…

- Flexible coupling infrastructure
- Memory scalability of all components
  - Minimize global arrays
- Performance scalability of all components
  - Hybrid MPI and OpenMP for multicore architecture
  - ALL active components - CAM, CLM, CICE and POP2 - now meet this requirement
- Parallel I/O throughout system
- Scalable Dynamical Core Option (HOMME)
CESM1 “Hub and Spoke” Coupling Architecture

Atm -> Coupler
- Bottom level temperature, pressure, wind…
- Downward Shortwave (vis, nir)
- Precipitation
- Carbon and Dust fluxes

Coupler -> Atm
(merged from Ind, ice and ocn)
- Latent, sensible heat fluxes
- Surfaces Stresses
- Upward long wave
- Evaporative water flux
- Surface Albedos

Note: Glacier (glc) component is new with CESM1 and not benchmarked here.
CESM: Coupler-7 Architecture
Climate Model Nomenclature/Details

- CCSM renamed CESM1 in June 2010
- Configuration Nomenclature - \( N^\circ \times M^\circ \)
  - \( N^\circ \) Atmosphere/Land models
  - \( M^\circ \) Ocean/Sea Ice models
- Example: component grids – (0.5° x 0.1°)
  - 0.50° ATM [576 x 384 x 26] - CAM
  - 0.50° LND [576 x 384 x 17] - CLM
  - 0.1° OCN [3600 x 2400 x 42] – POP2
  - 0.1° ICE [3600 x 2400 x 20] – CICE
- SYPD – simulated years per day
  - Throughput measure of performance
  - You won’t see flops mentioned here
CESM: Parallel I/O (PIO) Library

Rearranges data from model decomp to I/O friendly decomp

Interface between the model and the I/O library. Supports
- Binary
- NetCDF3 (serial netcdf)
- Parallel NetCDF (pnetcdf) (MPI/IO)
- NetCDF4

Big Global Array on process 0
How do we Load Balance Multi-component models?

Optimize throughput and decrease idle cycles

1. Increase core count for POP from 1664 cores to 3136 cores, reducing idle time from 1.53 SYPD to 2.23 SYPD.

2. Courtesy of John Dennis
CESM1 $0.5^\circ \times 0.1^\circ$ Scalability on Cray XT systems

1.9 years/day on 5844 cores with I/O on kraken hex-core XT5 (no threading)

(Courtesy of John Dennis)
CESM1 0.5° x 0.1°
Tuning Computational Efficiency

Cost CPU hours/year vs. Physical Processors

Sweet spot

- franklin-XT4 [no I/O]
- kraken-XT4 [no I/O]
- kraken-XT5 [no I/O]
CESM1 OpenMP/MPI
Cray XT5 Hex-core Scalability
0.25° x 0.1° case

(Courtesy of Pat Worley)
HRC06 Production Run
0.5° x 0.1° Details

• 155 year control run
• ~18M CPU hours
• 5844 cores for 4-5 months
• ~100 TB of data generated
• 0.5 to 1 TB per wall clock day generated
CESM1 0.1° ocean/sea-ice visualization

Note: atmospheric physics has not been retuned to its higher 0.5° resolution

This mov file was removed from pdf version to reduce size.
CESM1 Scalable HOMME Dycore: High-Order Methods Modeling Environment

• Algorithmic Advantages of High Order Methods
  – $h$-$p$ element-based method on quadrilaterals ($N_e \times N_e$)
  – Exponential convergence in polynomial degree ($N$)

• Computational Advantages of High Order Methods
  – Naturally cache-blocked $N \times N$ computations
  – Nearest-neighbor communication between elements (explicit)
  – Well suited to parallel microprocessor systems
HOMME: Quasi-uniform Cube-Sphere Grid

- Sphere is decomposed into 6 identical regions using a central projection (Sadourny, 1972) with equiangular grid (Rancic et al., 1996).
- Avoids pole problems, quasi-uniform.
- Non-orthogonal curvilinear coordinate system with identical

Ne=16 Cube Sphere
Showing degree of non-uniformity
CCSM/HOMME Scalability
0.125° atm / 0.25° land / 0.1° ocean

Work of Mark Taylor, Jim Edwards and Brian Eaton

- CCSM times include ALL CCSM components (PIO use was critical)
- Scalability of the dynamical core is preserved by CAM and scalability of CAM is preserved by CCSM
- Scale out to 86000 cores (BGP) and get 3 SYPD (Jaguarpf)

Estimated required capability To reach 1 km

<table>
<thead>
<tr>
<th>Label</th>
<th>APE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORNL–PF</td>
<td>0.25°</td>
</tr>
<tr>
<td>ANL–BGP</td>
<td>0.25°</td>
</tr>
<tr>
<td>ORNL–PF</td>
<td>0.125°</td>
</tr>
<tr>
<td>ORNL–PF</td>
<td>CCSM 0.125°</td>
</tr>
<tr>
<td>ANL–BGP</td>
<td>CCSM 0.25°</td>
</tr>
</tbody>
</table>
Next Generation NCAR/CISL Infrastructure
(Planning for 2012)

- Science Gateways: RDA, ESG
- Data Transfer Services
- Remote Vis
- Partner Sites
- TeraGrid Sites
- 10Gb/40Gb/100Gb Ethernet
- High Bandwidth I/O Network: QDR InfiniBand, 10Gb Ethernet
- HPSS 100 PB
- Data Collections, Project Spaces, Scratch, Archive Interface
- Data Analysis Visualization Nodes
- Compute Cluster: >1 PFLOPS
- Storage Cluster: 15 PB, 150 GB/s
New Systems and Facilities
The NCAR Wyoming Supercomputing Center:
A Petascale Facility
Dedicated to the Atmospheric Sciences
NWSC: Timeline

- **Ground-breaking & Construction Start**: June 2010
- **Construction Completion**: August 2011
- **NWSC Commissioning Completion**: December 2011
- **NWSC Support of Scientific Supercomputing Begins**: June 2012
NWSC Fact Sheet

- Project cost: ~$70M
- HPC system install begins January 2012
- Two, 1,115 m² raised floor areas
  - Power density – 6727 W/m²
  - Rack power day 1 – 4.5 MW
- 221-595 m² archive space
- Total floor area: 15,885 m²
  - Main floor: 14,242 m²
  - Upper floor: 1,643 m²
- Facility/site expandable to 24 MW
  - With additional funding of course!
- Distance from NCAR: 170 km
NWSC: Power Efficiency

NWSC PUE target: 1.10
Acknowledgements

• NCAR:
  • D. Bailey
  • F. Bryan
  • T. Craig
  • B. Eaton
  • J. Edwards [IBM]
  • N. Hearn
  • K. Lindsay
  • N. Norton
  • M. Vertenstein

• COLA:
  • J. Kinter
  • C. Stan

• U. Miami
  • B. Kirtman

• U.C. Berkeley
  • W. Collins
  • K. Yelick (NERSC)

• U. Washington
  • C. Bitz

• NICS:
  • M. Fahey
  • P. Kovatch

• ANL:
  • R. Jacob
  • R. Loy

• LANL:
  • E. Hunke
  • P. Jones
  • M. Maltrud

• LLNL
  • D. Bader
  • D. Ivanova
  • J. McClean (Scripps)
  • A. Mirin

• ORNL:
  • P. Worley

• Grant Support:
  • DOE
    • DE-FC03-97ER62402 [SciDAC]
    • DE-PS02-07ER07-06 [SciDAC]
  • NSF
    • Cooperative Grant NSF01
    • OCI-0749206 [PetaApps]
    • CNS-0421498
    • CNS-0420873
    • CNS-0420985

• Computer Allocations:
  • TeraGrid TRAC @ NICS
  • DOE INCITE @ NERSC
  • LLNL Grand Challenge

• Thanks for Assistance:
  • Cray, NICS, and NERSC

and many more…
Thanks! Any Questions?
NCAR/CISL HPC-Data Infrastructure
(Summer 2010)

Data Analysis Cluster

Visualization Cluster

Science Gateways
RDA, ESG

Data Transfer Services
(E.g., Gridftp)

Storage Cluster
"GLADE"
2 PB

I/O Network
10Gb

IBM
Power6
"Bluefire"
77 Tflops

IBM BlueGeneL
"Frost"
23 Tflops

Cray XT5
"Lynx"
8 Tflops

HPSS / MSS *
25 PB- Capacity
11 PB- Stored

3.8 GB/s
5.8 GB/s

* Migrating from NCAR Mass Storage System to HPSS during 2010
# CESM History File Sizes (GB)

<table>
<thead>
<tr>
<th></th>
<th>IPCC 1x1</th>
<th>0.5 x 0.1</th>
<th>0.25 x 0.1</th>
<th>0.125 x 0.1</th>
<th>Exascale 1 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>0.2</td>
<td>0.9</td>
<td>3.6</td>
<td>14.4</td>
<td>2250</td>
</tr>
<tr>
<td>LND</td>
<td>0.1</td>
<td>0.2</td>
<td>0.9</td>
<td>3.6</td>
<td>560</td>
</tr>
<tr>
<td>ICE</td>
<td>0.7</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>440</td>
</tr>
<tr>
<td>OCN</td>
<td>1.2</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>1900</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2.2</strong></td>
<td><strong>24.5</strong></td>
<td><strong>28.9</strong></td>
<td><strong>41.4</strong></td>
<td><strong>5150</strong></td>
</tr>
</tbody>
</table>