Computational & Information Systems Laboratory

The accumulated impact of code optimizations on scientific throughput of CESM on Cheyenne

John Dennis

Rory Kelly, Jim Edwards, Brian Dobbins, Chris Kerr, Youngsung Kim, Raghu Kumar, Sheri Mickelson, Rich Loft, Mariana Vertenstein, Sean Santos

Feb 21, 2018

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Overview

- Broadwell versus Knights Landing: A tail of different memory hierarchies
- Cost of Aquaplanet @ 100 km for several different dynamical cores
- The accumulated impact of code optimizations on scientific throughput of CESM on Cheyenne

Broadwell versus Knights Landing:
A tail of different memory hierarchies

Computational & Information Systems Laboratory

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

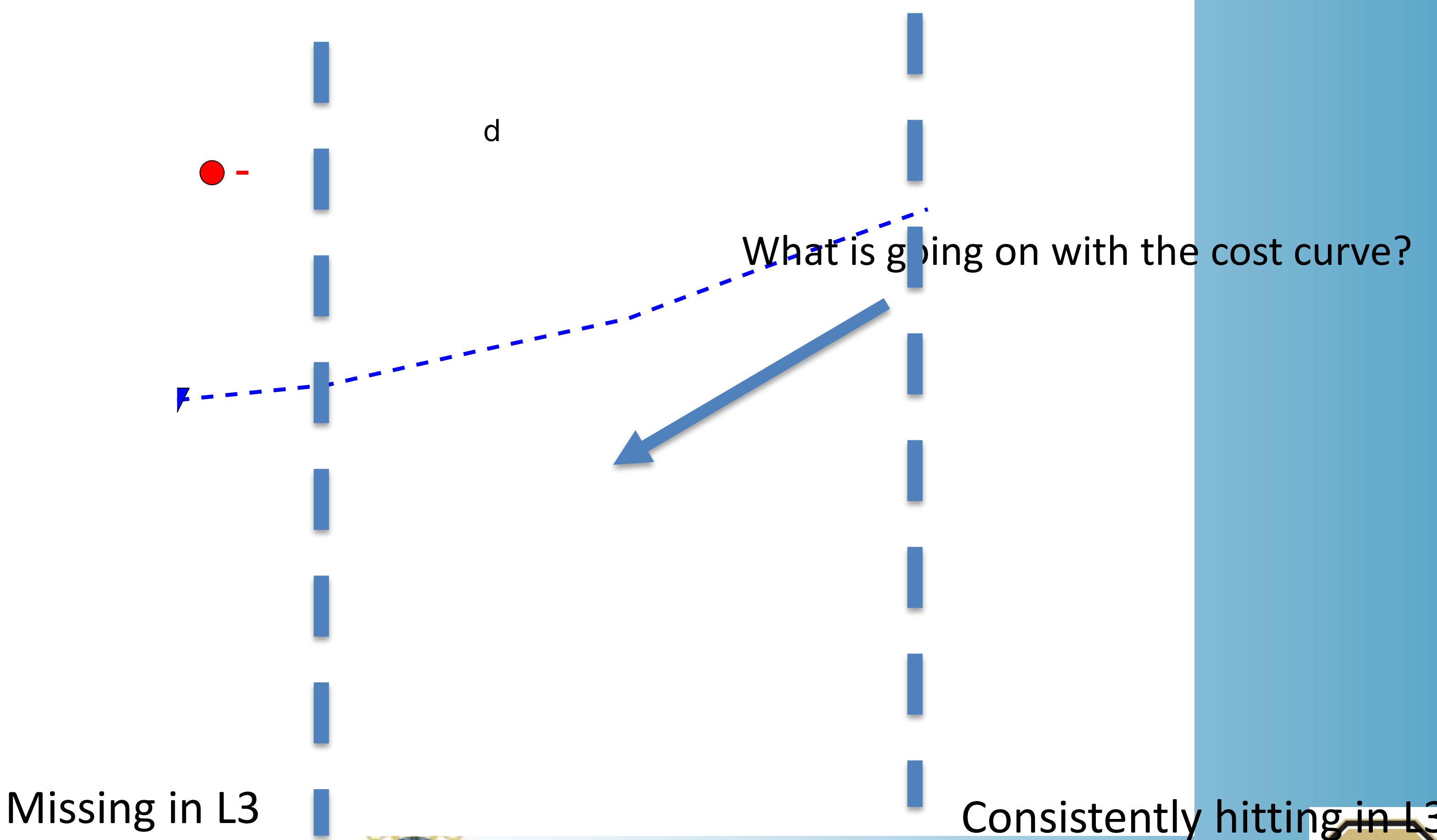
Related/Collaborative Activities

- Funding from Intel Parallel Computing Center (IPCC-WACS)
- NESAP (NERSC Exascale Science Application Program)
  – Bi-weekly: NERSC-Cray-NCAR telecon on CESM & HOMME performance (Feb 2015)
- Weekly Intel-TACC-NREL-NERSC-NCAR telecon
  – Concall focused on CESM/HOMME KNC performance

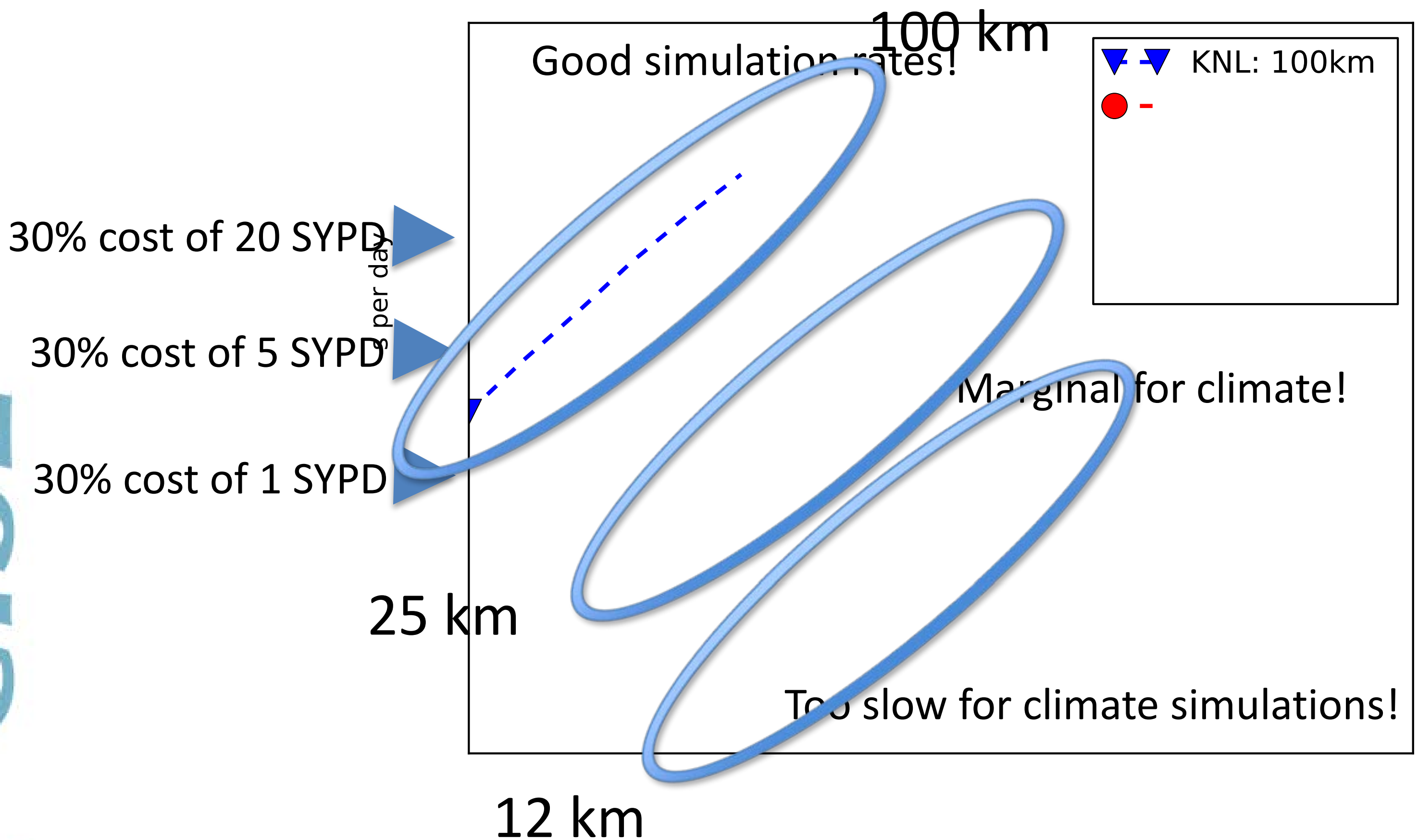
Optimization phases

1. Threading memory copy in boundary exchange [Jamroz]
2. Restructure data-structures for vectorization [Vadlamani & Dennis]
3. Rewrite message passing library/ specialized comm ops [Dennis]
4. Rearrange calculations in euler_step for cache reuse [Dennis]
5. Reduced # of divides [Dennis]
6. Restructured/alignment for better vectorization [Kerr]
7. Rewrote and optimized limiter [Demeshko & Kerr]
8. Redesign of OpenMP threading [Kerr & Dennis]
9. Flexible MPI message passing back-ends [Dennis]
   1. MPI_Put/Get (MPI3)
   2. MPI neighborhood collectives (MPI3)
10. Replaced all functions with subroutines [Kerr & Dennis]
11. Custom OpenMP barrier [Dobbins]

Simulation cost for HOMME on Xeon and Xeon Phi @ 100 km



Simulation rate for HOMME on Xeon and KNL



Cost of Aquaplanet @ 100 km for several different dynamical cores

Aquaplanet @ 100 km

Dynamical core + advection algorithm	MPI rank x OpenMP threads	Capability (sim yr/day)	Cost (core-hrs per sim yr)	Increase/decrease relative to CAM-fv @ 1152x3
CAM-fv	1152x3	40.4	2053	0.0%
CAM-SE/ eulerian	2700x1	33.6	1929	-6.0%
	5400x1	58.8	2203	7.3%
CAM-SE/ CSLAM	2700x1	29.8	2173	5.8%

Computational & Information Systems Laboratory

CISL

NCAR

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

NSF

Computational & Information Systems Laboratory

The accumulated impact of code optimizations on scientific throughput of CESM on Cheyenne

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Two related code optimization investments

- ASAP led effort
  - An effort to explore the use of accelerator and other future technology on existing weather and climate model codes
- Strategic Parallel Optimization and Optimization Computing (SPOC)

  An NCAR-wide effort to increase the performance and efficiency of NCAR community does on CESM, WRF, and MPAS

Approach: Incrementally improve existing codebase

Estimate impact

- What impact did this investment have on scientific throughput?
- Challenging because CESM code base has changed from both a scientific and code optimization perspective
- Approach
  - Detailed measurements of execution time of CESM2 on Cheyenne
  - Adjust execution time of segments of the code based historical timing information
    - I.e. reduced execution time of short-wave length radiation by 33% on Yellowstone….

Computational & Information Systems Laboratory

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

What was done?

- Optimized the following pieces of CESM
  - Aerosol wet deposition
  - Morrison Gettelman micro-physcs
  - Rapid radiative transport model
  - Planetary boundary layer
  - Heterogenous freezing in the Modal Aerosol Model
  - Random number generator
  - Implicit chemical solver
  - Spectral element dynamical core (SE-dycore)
  - CSLAM advection algorithm in the SE-dycore
  - CICE boundary exchange *
  - Better load balance of CESM

* Not currently reintegrated back into code base

What was achieved ?
Higher efficiency = more science

CESM configuration	Atmos Resolution (km)	Ocean Resolution (km)	Speedup
Low-res IPCC	100	100	13%
Low-res WACCM chemistry	100	100	20%
High-res IPCC	25	100	25%
Ultra-high Ocean eddy permitting	25	10	35%

What was achieved ?
Higher efficiency = more science

CESM configuration	Atmos Resolution (km)	Ocean Resolution (km)	Speedup
Low-res IPCC	100	100	13%
Low-res WACCM chemistry	100	100	20%
High-res IPCC	25	100	25%
Ultra-high Ocean eddy permitting	25	10	35%

- CCSM/CESM consumes 57% of all Cheyenne
- The TCO to provision 1% more climate computing is $285K over the 4-year life of Cheyenne
- Investment has enabled between $3.7M and $9.9M of additional science throughput on Cheyenne. Since CESM is a community model the valuation is larger.

What was achieved ?
Simulation rate @ 100 km

NCAR System	Intel Xeon Processor	CESM Version	Capability (sim yr/day)	Cost (core-hrs per sim yr)
Cheyenne	18c Broadwell	CESM2	30	3500
Yellowstone	8c Sandybridge	CESM2	19.6	5167
Yellowstone	8c Sandybridge	CESM1	10.6	1521

What was achieved ?
Simulation rate @ 100 km

NCAR System	Intel Xeon Processor	CESM Version	Capability (sim yr/day)	Cost (core-hrs per sim yr)
Cheyenne	18c Broadwell	CESM2	30	3500
Yellowstone	8c Sandybridge	CESM2	19.6	5167
Yellowstone	8c Sandybridge	CESM1	10.6	1521

- Cheyenne is 48% more efficient than Yellowstone for production configurations of CESM2
- CESM2 on Cheyenne can deliver 2.8x the capability compared to CESM1 on Yellowstone

Challenges

- New code being added 30x quicker then it can be optimized
  - CESM1: 1.3M lines of code
  - CESM2: 1.6M lines of code
  - ~10K lines of code has been optimized
- Scientific evolution of codebase is unpredictable
  - SE dynamical core
  - Cloud Layer Unified by Bi-normals (CLUBB)
- Use of Github should greatly simplify reintegration effort.
- Most optimization efforts performed twice

Conclusions/Future work

- Concerted/sustained effort reduced cost of HOMME on Xeon and Xeon Phi
  - 2x speedup on KNL and BDW
  - Optimizations reduces time to solution by 32x for similar simulation cost
- Modest difference in cost of CAM using different dynamical cores at production core counts
- Investment in code optimization increased scientific throughput of Cheyenne by \$3.7M to \$9.9M
- Improves in process efficiency can likely reduce ratio of added to optimized code to 10x
  - Optimizing code after insertion not a viable long term approach
  - Teaching code optimization: RRTMGP (Next generation radiation model) & Robert Pincus
- Incremental approach does not address transformative architecture changes

Questions?

dennis@ucar.edu

Computational & Information Systems Laboratory

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Motivation of HOMME optimization effort

- Atmosphere dynamical core (HOMME)
  - CAM: 35% of time (vert levels=32, # of tracers=25)
- Much easier to optimize then physics ☺
- Benchmark code
  - CORAL (CAM-SE)
  - NSF625
- Useful for evaluating full system performance

Group/Team

- Rich Loft, Division Director (NCAR)
- John Dennis, Scientist (NCAR)
- Chris Kerr, Software Engineer, contractor
- Youngsung Kim, Software Engineer (NCAR) /Graduate Student (CU)
- Brian Dobbins, Software Engineer (NCAR)
- Raghu Raj Prasanna Kumar, Associate Scientist (NCAR)
- Sheri Mickelson, Software Engineer (NCAR) / Graduate Student (CSU)
- Ravi Nanjundiah, Professor (IISc)

Energy usage for HOMME (NGGPS-like) on Xeon and Xeon Phi @ 12 km