Performance Results for the new CAM Benchmark Suite

Patrick Worley
Oak Ridge National Laboratory

Arthur Mirin
Lawrence Livermore National Laboratory

13th Annual CCSM Workshop
June 17-19, 2008
Breckenridge, CO

Acknowledgements

- The work described in this presentation was sponsored by the Atmospheric and Climate Research Division and by the Office of Mathematical, Information, and Computational Sciences, both of the Office of Science, U.S. Department of Energy, under Contract No. DE-AC05-00OR22725 with UT-Battelle, LLC and Contract No. DE-AC52-07NA27344 with Lawrence Livermore National Security, LLC.
- These slides have been authored by contractors of the U.S. Government under contracts No. DE-AC05-00OR22725 and No. DE -AC52-07NA27344, and are released as LLNL Report LLNL -PRES-404700. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.
- This work used resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the Department of Energy under Contract DE -AC05-00OR22725.

Goals

- Goals of Benchmarking Activity
  - Quantify change in cost of model due to model evolution and to choice of model configuration.
  - Establish new performance baselines upon which to base future performance optimization activities, and with which to evaluate computing platforms.
- Goals of Presentation
  - Outline benchmark suite and describe recent performance results.

Participants

- Benchmark specification
  - Phil Rasch (lead)
  - Philip Cameron-Smith
  - Brian Eaton
  - Cecile Hannay
  - Peter Hess
  - Jean-Francois Lamarque
  - Art Mirin
  - Will Sawyer
  - Pat Worley
- Initial benchmarking
  - Pat Worley

Benchmark Suite

- cam3.0p1

  default configuration (C0): T42L26, T85L26 (spectral Eulerian dynamical core)

- cam3.1p2

  C0: FV1.9x2.5 L26 (Finite Volume dycore), T42L26, T85L26

- CAM from development trunk (cam3_5_27 and later)

  C0: FV1.9x2.5 L26, T42L26, T85L26

  C0r: C0 with RRTMG radiation package (FV1.9x2.5 L26 only; cam3_5_42)

  C1: C0 with 30 levels and FV1.9x2.5 only

  C2: C1 with “cam3.5” aerosols

Benchmark Suite

- CAM from development trunk

  C3: C2 with UW physics package

  C4: C3 with Morrison Gettelman cloud parameterization

  C5: C4 with predicted aerosol fields

  C6: C4 with full tropospheric chemistry (cam3_5_45)

Benchmark Suite Comments

- cam3.0p1 => cam3.1p2:
  - new physics interface design (software engineering release)
- cam3.1p2 => cam3_5_27:
  - Neale-Richter convection mods, Vavrus “freeze-dry” mod, surface component coupling every timestep, …
- C0 => C1: 30 levels
- C1 => C2
  - Prescribed aerosol datasets generated by CAM with prognostic aerosol chemistry turned on; reimplementation of existing interpolation scheme

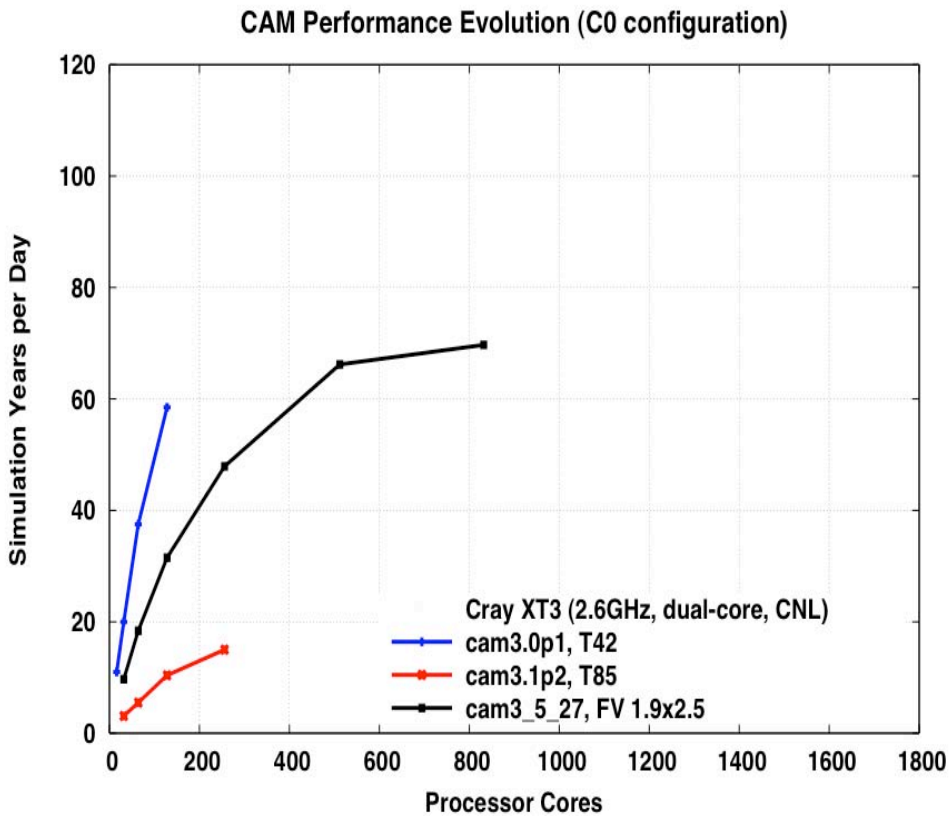
Benchmark Suite Comments

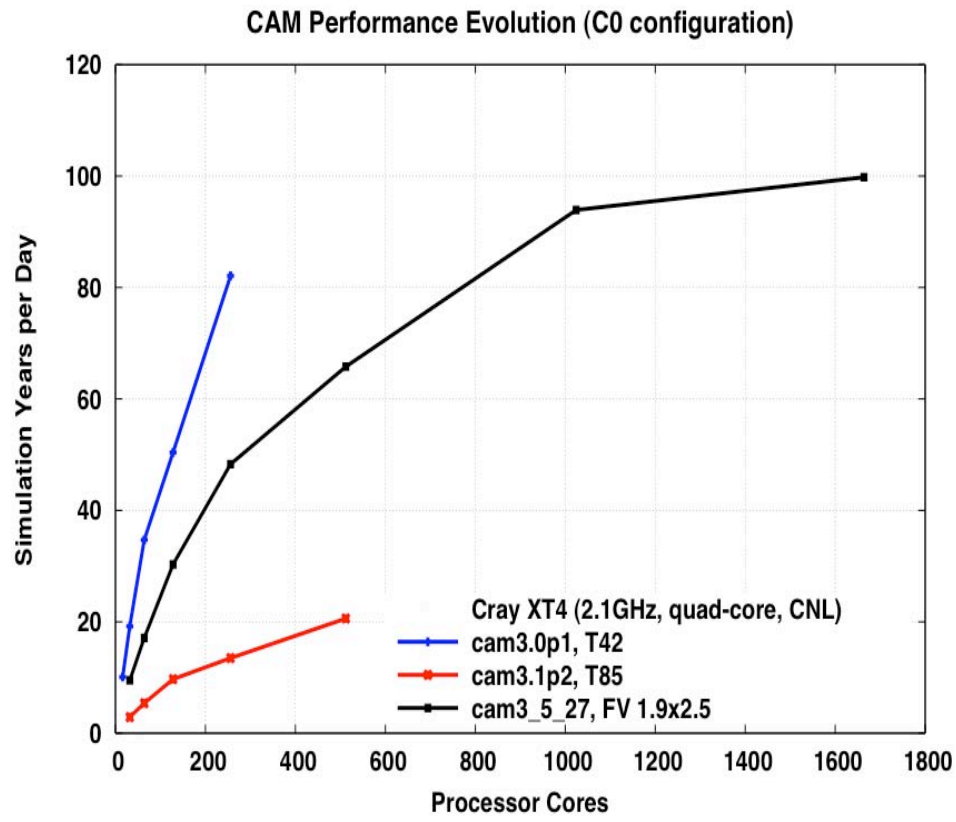
- C2 => C3
  - Replace Holtslag-Boville vertical diffusion with diag_TKE (Grenier-Bretherton); replace Hack shallow convection with UW (McCaa)
- C3 => C4
  - Replace Rasch-Kristjansson microphysics with Morrison Gettelman, including addition of two advected species
- C4 => C5
  - Enable prognostic aerosol code, adding 16 advected species
- C4 => C6
  - Enable full tropospheric chemistry, adding 103 advected species

Benchmark Details

- 30 simulation days (September) with monthly history output
- Platforms (all supporting MPI and OpenMP):
  - Cray XT3 at Oak Ridge National Laboratory (ORNL): dual -core 2.6 GHz Opteron processors running Compute Node Linux (CNL)
  - Cray XT4 at ORNL: quad-core 2.1 GHz Opteron processors running CNL
  - IBM BG/P at ORNL: quad-core 850MHz PowerPC processors
- Set optimization as high as possible while still preserving reproducibility. CAM has not been validated on any of these platform as of yet.
- Experiments examined both pure MPI and hybrid MPI/OpenMP.

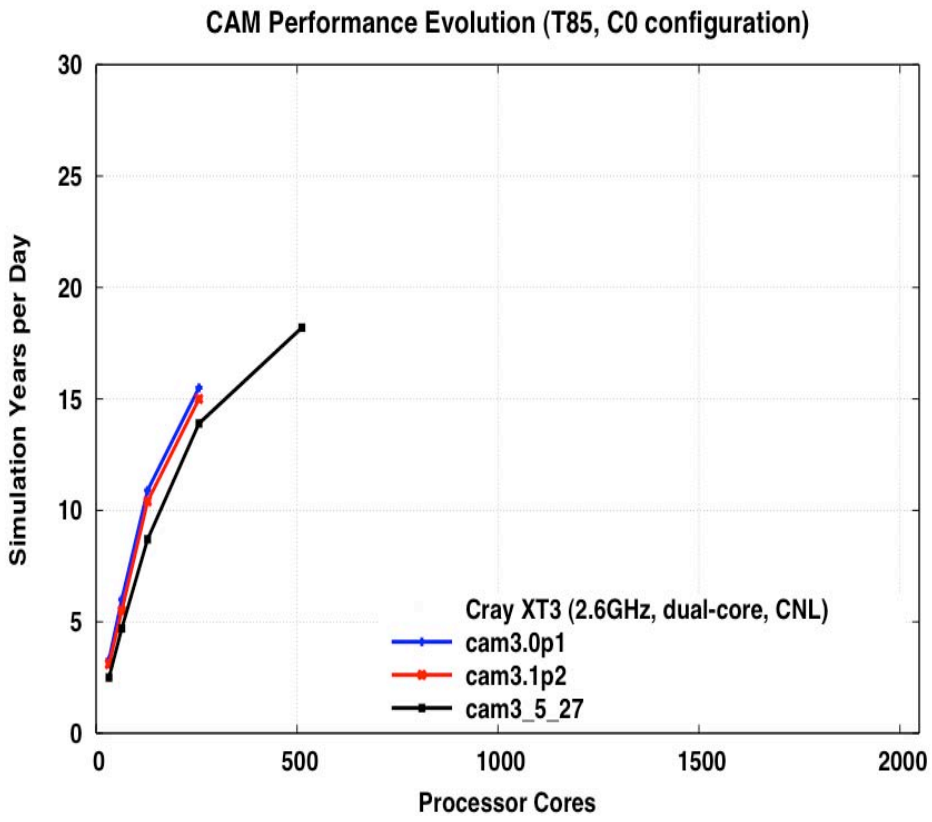
3.0 (T42) vs. 3.1 (T85) vs. 3.5 (FV1.9)

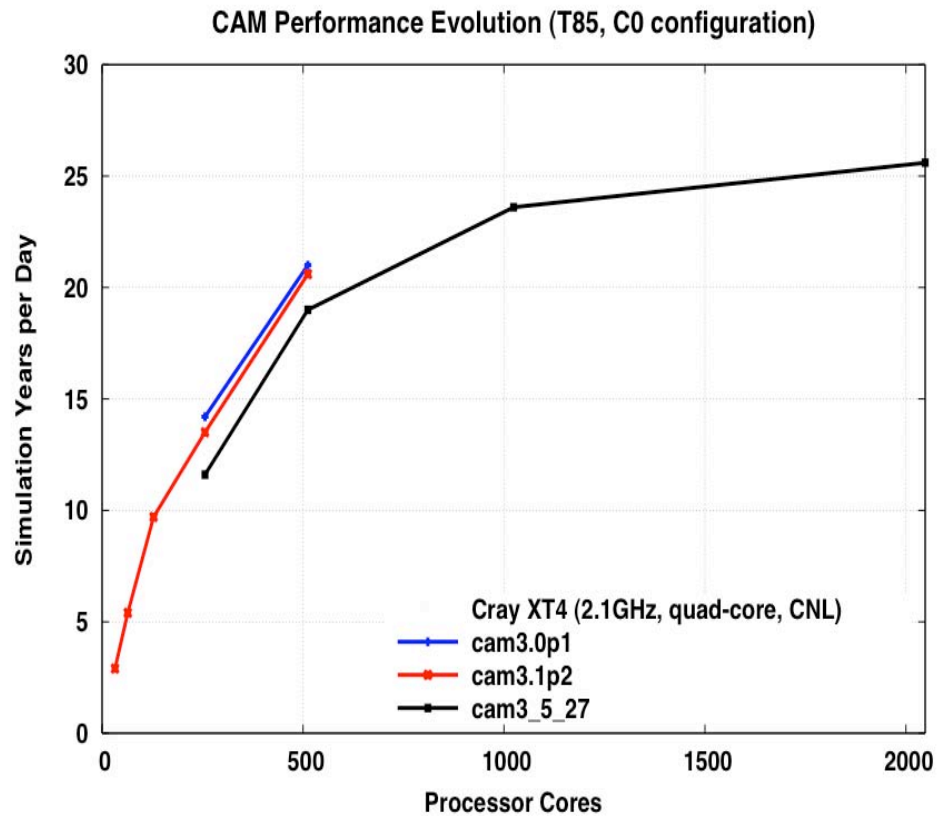




Analysis compares performance of recent CAM production configurations. (cam3.5 can be run with even higher processor core counts, but performance does not continue to increase on these platforms for this benchmark.)

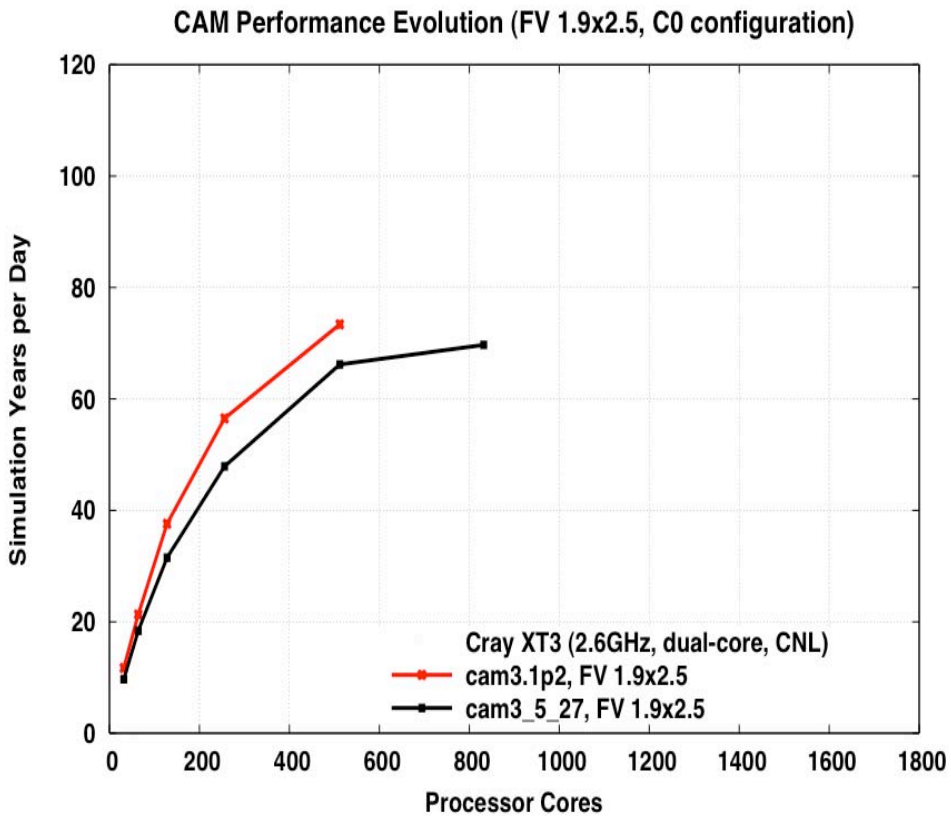
T85: 3.0 vs. 3.1 vs. 3.5

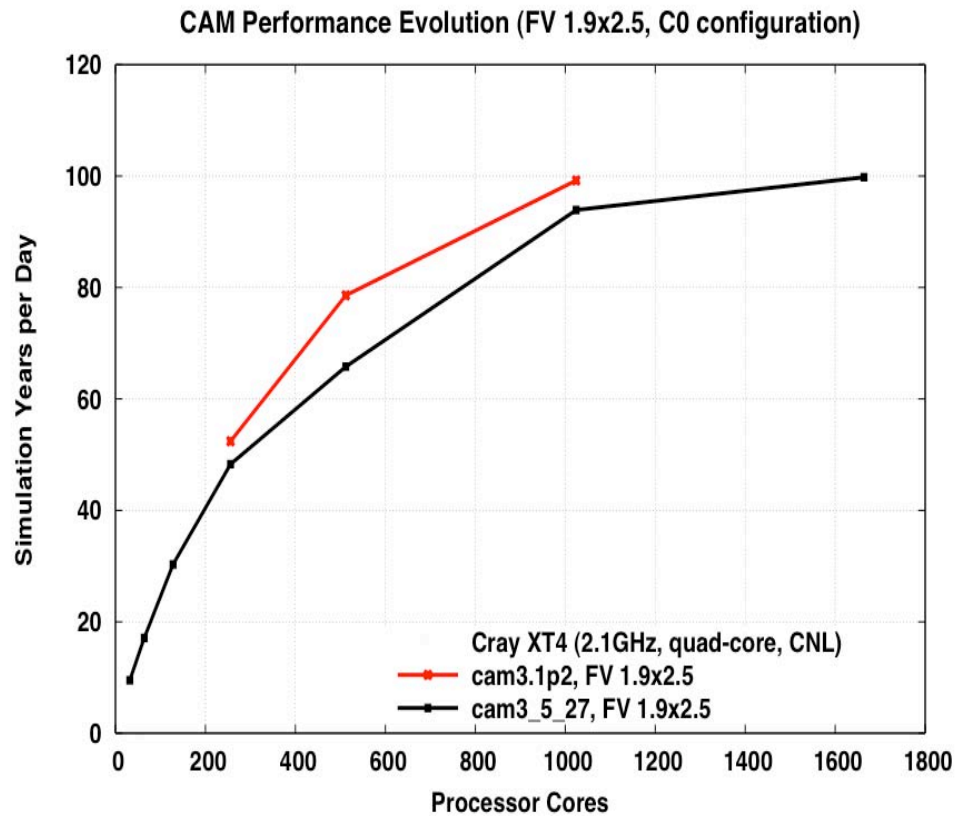




For small processor counts, cam3.5 is 20-30% slower than cam3.0 . However, cam3.5 is also more scalable than the released version of cam3.0 or cam3.1 . Data still to be collected for 1024 cores on XT3 and for small core counts on XT4.

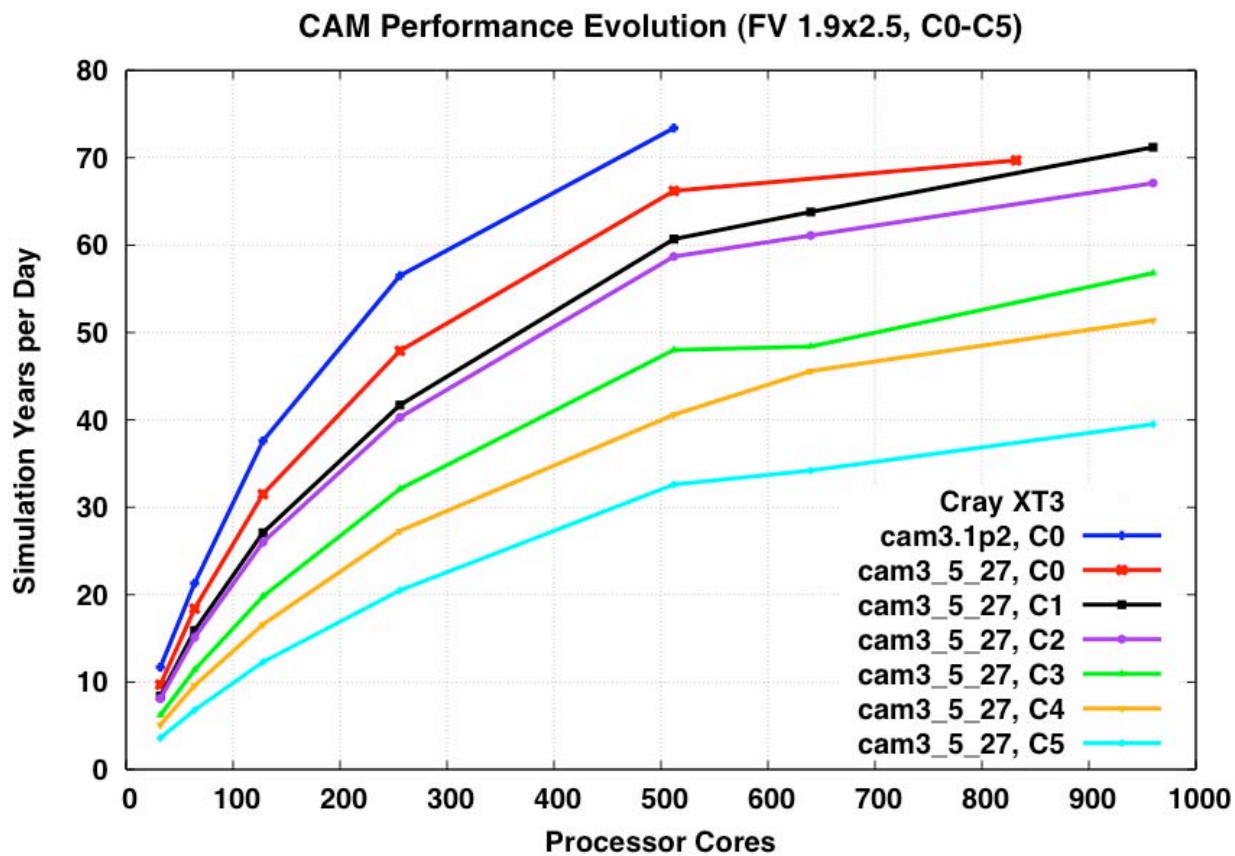
FV: 3.0 vs. 3.1 vs. 3.5





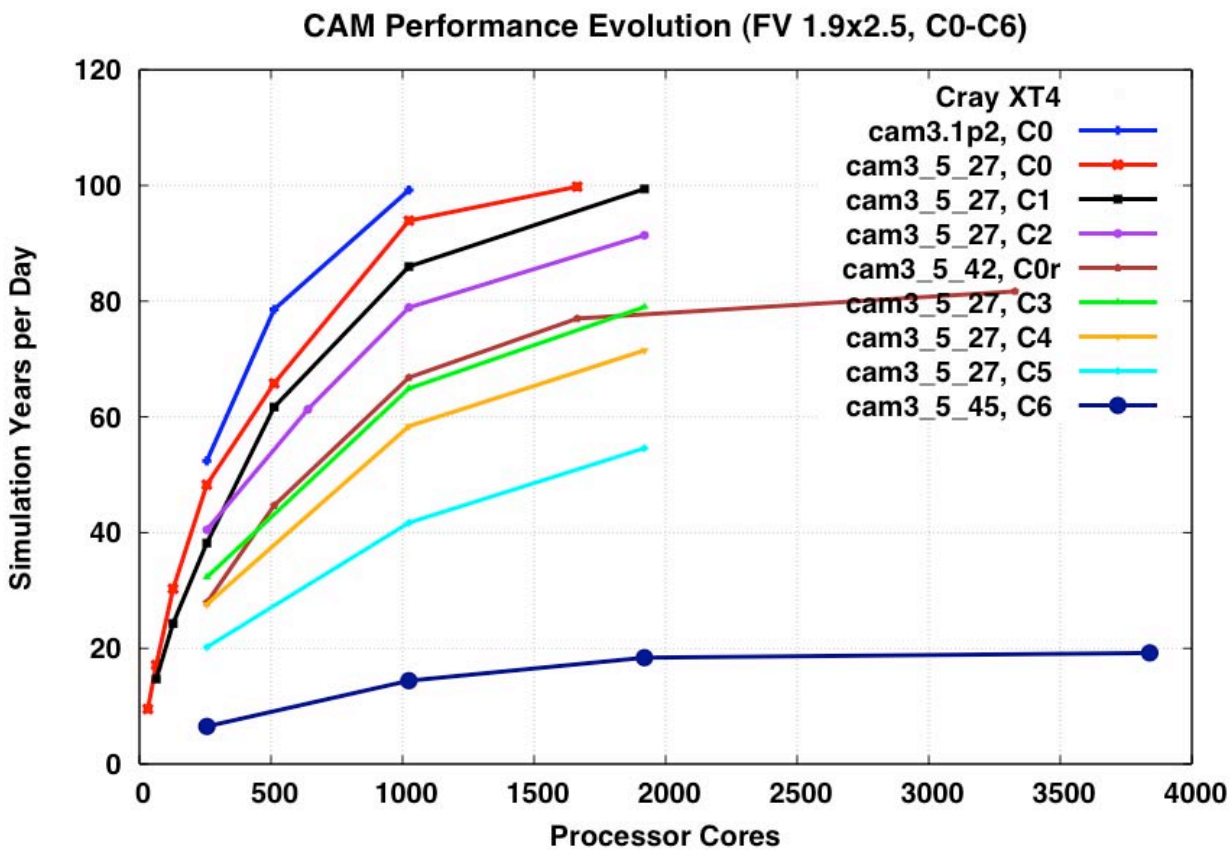
For small processor counts, cam3.5 is 20% slower than cam3.1 .
Additional scalability does not recover lost performance (currently).

XT3 FV: C0 vs. C1 vs. C2 vs. C3 vs. C4 vs. C5



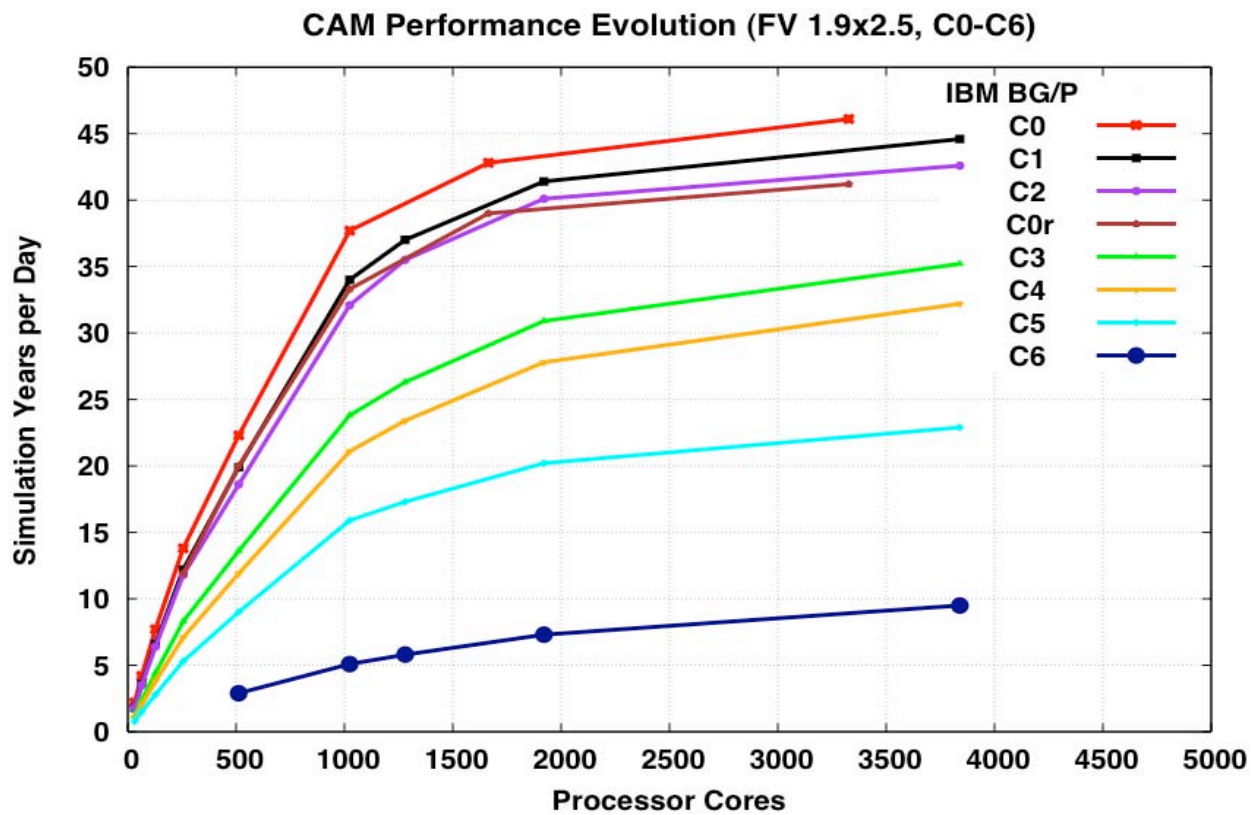
For 32 and 512 processor cores, the normalized cost progression is
.83 => 1.00 => 1.15 => 1.20 => 1.56 => 1.90 => 2.69 and
.90 => 1.00 => 1.09 => 1.13 => 1.38 => 1.63 => 2.03, respectively.

XT4 FV: C0 vs. C1 vs. … vs. C6



For 1024 processor cores, the normalized cost progression (C0 through C6) is .95 => 1.00 => 1.09 => 1.19 => 1.45 => 1.61 => 2.25 => 6.52. C0r is 1.41 times as expensive as C0.

BGP FV: C0 vs. C1 vs. … vs. C6



For 1024 processor cores, the normalized cost progression (C0 through C6) is 1.00 => 1.13 => 1.20 => 1.60 => 1.80 => 2.37 => 7.39, and C0r is 1.13 times as expensive as C0.

Discussion

- It is worth using 1024 processors even for “small” 1.9x2.5 degree problems on quad-core systems when using OpenMP and MPI.
- Cost increase from C0 => C5 is less than a factor of 3.
- C6 benchmark is 7 times more expensive than C0. Alternatives exist that provide similar accuracy but at less cost (using fewer tracers). We will examine this in the future.
- Relative cost of RRTMG is significantly different between XT4 and BG/P systems. Ongoing performance analysis and optimization of RRTMG may shed light on this peculiarity.
- Performance scaling is reasonable, but there are many optimization options (see Mirin/Worley poster) and it takes time to determine the optimal performance for any given configuration. Performance may be somewhat greater than reported here.

Discussion

- A number of poorly scaling algorithms have been identified as part of this exercise. Some have been replaced, but others still need to be optimized or replaced. This will improve both performance and performance scalability.
- CAM performance scalability is improved by running in hybrid MPI/OpenMP mode on the Cray XT and IBM BG/P systems. We expect the same to be true for the other components of the CCSM.
- More data needs to be collected on the Cray XT4.
- As a function of node count, the quad-core XT4 system is faster than dual-core XT3 system in these experiments, despite the lower clock speed. The XT4 is 3 times faster than the BG/P for the same core count and 2 times faster for maximum achieved throughput rate.

What's Next?

- Continue benchmarking, especially on Cray XT4 and on IBM Power5/6 systems at NCAR. Repeat selected experiments on all platforms as model evolves, especially after parallel algorithm optimizations.
- Reduce number of benchmark configurations after October freeze.
- Add high resolution configuration(s):
  - FV 0.47x0.63? Others?
  - How many vertical levels? Should we add vertical resolution scaling studies?
  - WACCM benchmark?
- Add an aquaplanet benchmark (something suitable for initial comparison with alternative dycores)?
- Use data to guide optimization activities.