Next generation climate modelling

Tim Palmer
Department of Physics
University of Oxford

THE ROYAL SOCIETY

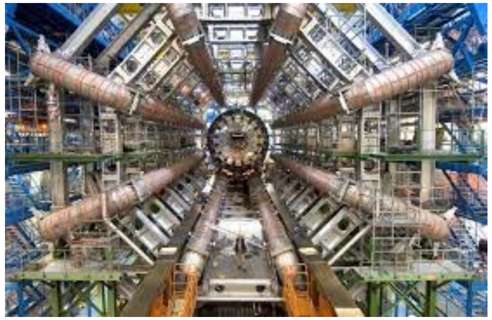
European Research Council
Established by the European Commission

Oxford Physics

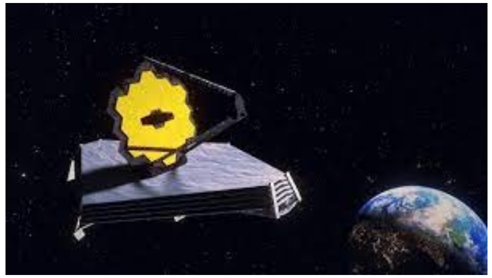
Dear WCRP members

Please, please, please think big!!

- Large Hadron Collider ($10 billion)



- James Webb Telescope ($10 billion)



Is understanding/predicting our climate really so much less important than understanding/predicting the world on quantum or cosmological scales?

We will shortly be in the era of the exascale (already here for some 16-bit arithmetic).

What are we planning to do with dedicated exascale computers?

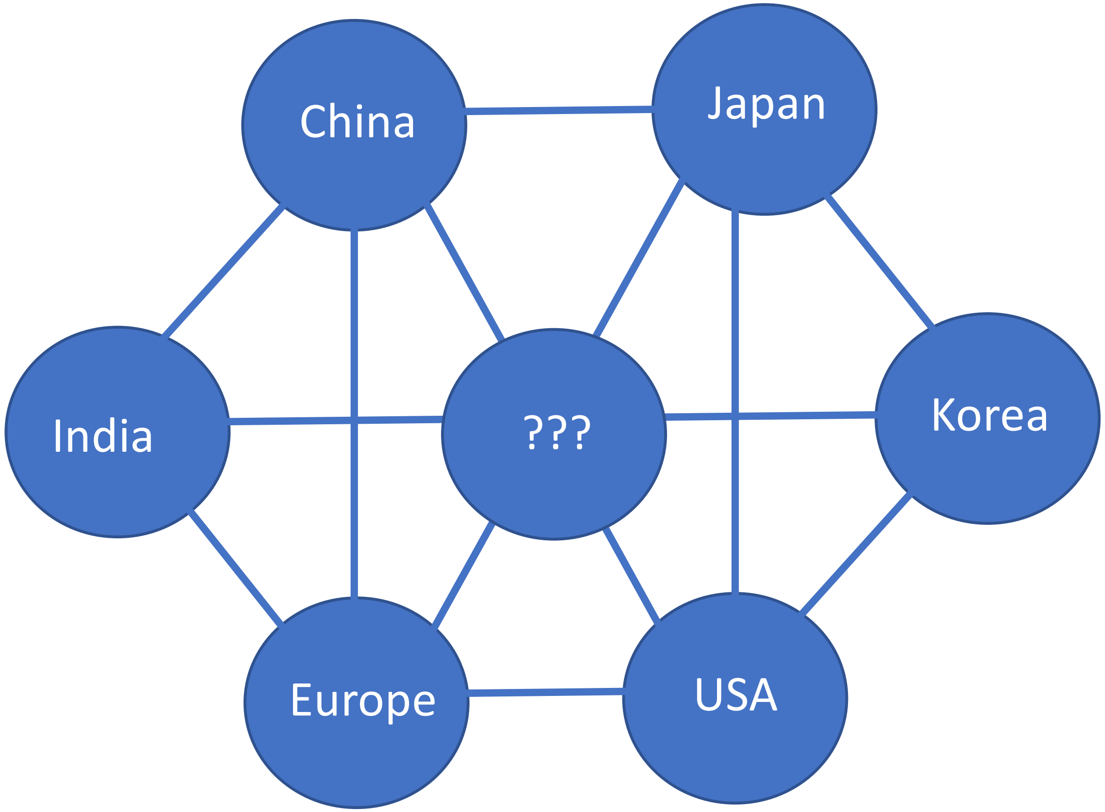
DOE/NNSA's El Capitan reaches 2 exaflops - more powerful than the top 200 fastest supercomputers in the world combined - to support nation's nuclear security missions

SAN JOSE, Calif. – March 4, 2020 – Hewlett Packard Enterprise (HPE) today announced that it will deliver the world's fastest exascale-class supercomputer for the U.S. Department of Energy's (DOE) National Nuclear Security Administration (NNSA) at a record-breaking speed of 2 exaflops - 10X faster than today's most powerful supercomputer. The new system, which the DOE's Lawrence Livermore National Laboratory (LLNL) has named El Capitan, is expected to be delivered in early 2023 and will be managed and hosted by LLNL for use by the three NNSA national laboratories: LLNL, Sandia National Laboratories and Los Alamos National Laboratory. The system will enable advanced simulation and modeling to support the U.S. nuclear stockpile and ensure its reliability and security.

Hewlett Packard Enterprise

Are we content with that?

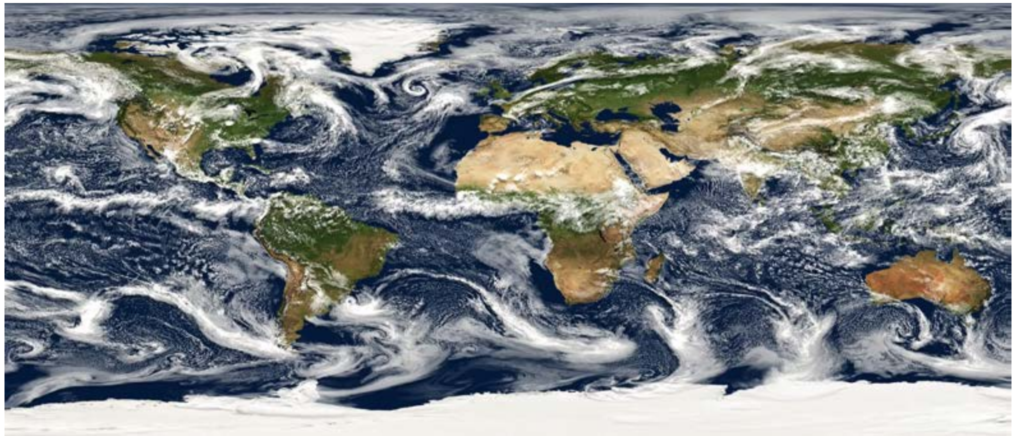
What could we achieve if a) we all worked together and b) persuaded our governments we need dedicated exascale machines to make reliable regional climate predictions/projections?

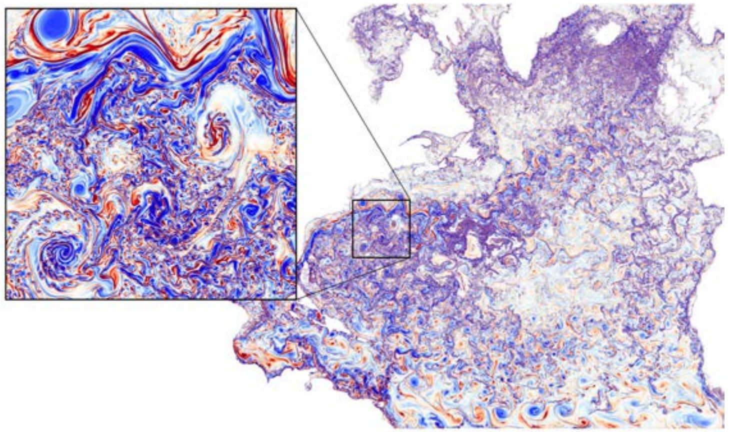


A coordinated international network of climate-dedicated exascale computing institutes.

And what would we do with such a distributed facility?

Coordinated km-scale (K-scale) global ensembles for multi-decadal climate prediction





Why km-scale (K-scale)?

- More accurate representation of underlying laws of physics.
- No parametrization of deep convection, orographic gravity wave drag and ocean eddy mixing.
- Smaller systematic errors.
- Better representation of extremes (vital for adaptation applications).
- Better assimilation of observations (ocean and atmosphere) and hence more accurate initial conditions.
- Better reanalysis products for climate diagnostics.
- Improved (stochastic) parametrisations for CMIP-class models.

PERSPECTIVE

PERSPECTIVE

The scientific challenge of understanding and estimating climate change

Tim Palmer[a,1] and Bjorn Stevens[b,1]

Edited by Jagadish Shukla, George Mason University, Fairfax, VA, and accepted by Editorial Board Member Robert E. Dickinson October 21, 2019 (received for review May 21, 2019)

Given the slow unfolding of what may become catastrophic changes to Earth's climate, many are understandably distraught by failures of public policy to rise to the magnitude of the challenge. Few in the science community would think to question the scientific response to the unfolding changes. However, is the science community continuing to do its part to the best of its ability? In the domains where we can have the greatest influence, is the scientific community articulating a vision commensurate with the challenges posed by climate change? We think not.

climate change | climate models | high resolution

The idea that the science of climate change is largely "settled," common among policy makers and environmentalists but not among the climate science commu- how this warming plays out regionally, and what it implies for the likelihood of surprises. In our view, the political situation, whereby some influential

High resolution, ensembles, earth-system complexity. Incompatible? No!

- Can significantly reduce computational cost of increased resolution by running climate models with mixed precision (32 - 16-bit) floating point arithmetic (up to x32 according to NVIDIA). Cost of earth-system complexity could be reduced even more with ultra-fast neural nets trained on off-line earth-system modules.

15 February 2022 PAXTON ET AL. 1215

Climate Modeling in Low Precision: Effects of Both Deterministic and Stochastic Rounding

E. Adam Paxton,[a] Matthew Chantry,[a] Milan Klöwer,[a] Leo Saffin,[b] and Tim Palmer[a]

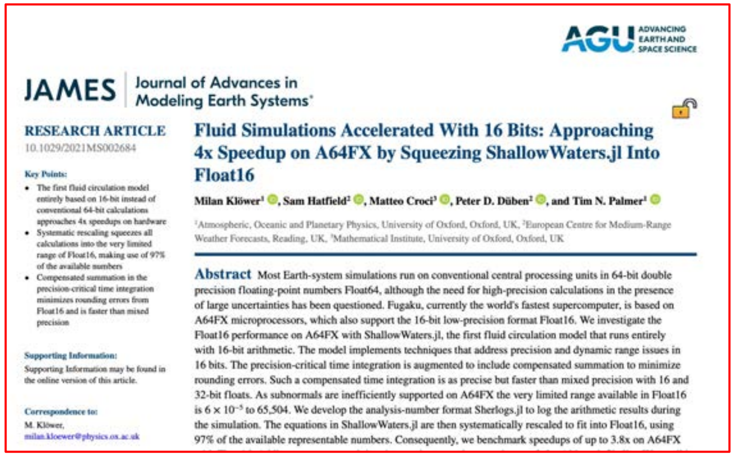
[a] *University of Oxford, Oxford, United Kingdom*
[b] *University of Leeds, Leeds, United Kingdom*

(Manuscript received 4 May 2021, in final form 17 October 2021)

ABSTRACT: Motivated by recent advances in operational weather forecasting, we study the efficacy of low-precision arithmetic for climate simulations. We develop a framework to measure rounding error in a climate model, which provides a stress test for a low-precision version of the model, and we apply our method to a variety of models including the Lorenz system, a shallow water approximation for flow over a ridge, and a coarse-resolution spectral global atmospheric model with simplified parameterizations (SPEEDY). Although double precision [52 significant bits (sbits)] is standard across operational climate models, in our experiments we find that single precision (23 sbits) is more than enough and that as low as half precision (10 sbits) is often sufficient. For example, SPEEDY can be run with 12 sbits across the code with negligible rounding error, and with 10 sbits if minor errors are accepted, amounting to less than 0.1 mm $(6\ h)^{-1}$ for average gridpoint precipitation, for example. Our test is based on the Wasserstein metric and this provides stringent nonparametric bounds on rounding error accounting for annual means as well as extreme weather events. In addition, by testing models using both round-to-nearest (RN) and stochastic rounding (SR) we find that SR can mitigate rounding error across a range of applications, and thus our results also provide some evidence that SR could be relevant to next-generation climate models. Further research is needed to test if our results can be generalized to higher resolutions and alternative numerical schemes. However, the results open a promising avenue toward the use of low-precision hardware for improved climate modeling.

SIGNIFICANCE STATEMENT: Weather and climate models provide vital information for decision-making, and will become ever more important in the future with a changed climate and more extreme weather. A central limitation to improved models are computational resources, which is why some weather forecasters have recently shifted from conventional 64-bit to more efficient 32-bit computations, which can provide equally accurate forecasts. Climate models, however, still compute in 64 bits, and adapting to lower precision requires a detailed analysis of rounding errors. We develop methods to quantify rounding error in a climate model, and find similar precision acceptable across weather and climate models, with even 16 bits often sufficient for an accurate climate. This opens a promising avenue for computational efficiency gains in climate modeling.

KEYWORDS: Error analysis; Numerical analysis/modeling; Climate models



AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2021MS002684

Fluid Simulations Accelerated With 16 Bits: Approaching 4x Speedup on A64FX by Squeezing ShallowWaters.jl Into Float16

Milan Klöwer[1], Sam Hatfield[2], Matteo Croci[3], Peter D. Düben[2], and Tim N. Palmer[1]

[1]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK, [2]European Centre for Medium-Range Weather Forecasts, Reading, UK, [3]Mathematical Institute, University of Oxford, Oxford, UK

Key Points:
- The first fluid circulation model entirely based on 16-bit instead of conventional 64-bit calculations approaches 4x speedups on hardware
- Systematic rescaling squeezes all calculations into the very limited range of Float16, making use of 97% of the available numbers
- Compensated summation in the precision-critical time integration minimizes rounding errors from Float16 and is faster than mixed precision

Supporting Information:
Supporting Information may be found in the online version of this article.

Correspondence to:
M. Klöwer,
milan.kloewer@physics.ox.ac.uk

Abstract Most Earth-system simulations run on conventional central processing units in 64-bit double precision floating-point numbers Float64, although the need for high-precision calculations in the presence of large uncertainties has been questioned. Fugaku, currently the world's fastest supercomputer, is based on A64FX microprocessors, which also support the 16-bit low-precision format Float16. We investigate the Float16 performance on A64FX with ShallowWaters.jl, the first fluid circulation model that runs entirely with 16-bit arithmetic. The model implements techniques that address precision and dynamic range issues in 16 bits. The precision-critical time integration is augmented to include compensated summation to minimize rounding errors. Such a compensated time integration is as precise but faster than mixed precision with 16 and 32-bit floats. As subnormals are inefficiently supported on A64FX the very limited range available in Float16 is 6×10^{-5} to 65,504. We develop the analysis-number format Sherlogs.jl to log the arithmetic results during the simulation. The equations in ShallowWaters.jl are then systematically rescaled to fit into Float16, using 97% of the available representable numbers. Consequently, we benchmark speedups of up to 3.8x on A64FX

- Dedicated exascale computing will allow small ensembles (<10 members) of K-scale multi-decadal integrations

- A federated network of exascale machines will allow a coordinated multi-model ensemble (>50 members) of K-scale integrations (c.f. DEMETER, ENSEMBLES, CMIP).

It can be done.

For the sake of society, it must be done.

With endorsement by WCRP and active engagement by WCRP members in the coming years, it will be done!

Providing we are prepared to think big!