



A simulation of Earth's atmosphere generated by the Community Atmosphere Model.

Build imprecise supercomputers

Energy-optimized hybrid computers with a range of processor accuracies will advance modelling in fields from climate change to neuroscience, says **Tim Palmer**.

Today's supercomputers lack the power to model accurately many aspects of the real world, from the impact of cloud systems on Earth's climate to the processing ability of the human brain. Rather than wait decades for sufficiently powerful supercomputers — with their potentially unsustainable energy demands — it is time for researchers to reconsider the basic concept of the computer. We must move beyond the idea of a computer as a fast but otherwise traditional 'Turing machine', churning through calculations bit by bit in a sequential, precise and reproducible manner.

In particular, we should question whether all scientific computations need to be performed deterministically — that is, always producing the same output given the same

input — and with the same high level of precision. I argue that for many applications they do not.

Energy-efficient hybrid supercomputers with a range of processor accuracies need to be developed. These would combine conventional energy-intensive processors with low-energy, non-deterministic processors, able to analyse data at variable levels of precision. The demand for such machines could be substantial, across diverse sectors of the scientific community.

MORE WITH LESS

Take climate change, for example. Estimates of Earth's future climate are based on solving nonlinear (partial differential) equations for fluid flow in the atmosphere and oceans. Current climate simulators — typically with

grid cells of 100 kilometres in width — can resolve the large, low-pressure weather systems typical of mid-latitudes, but not individual clouds. Yet modelling cloud systems accurately is crucial for reliable estimates of the impact of anthropogenic emissions on global temperature¹.

The resolution of this computational grid is determined by the available computing power. Current petaflop computers can perform up to 10^{15} additions or multiplications — floating-point operations — per second (flops). By the early 2020s, next-generation exaflop supercomputers, capable of 10^{18} operations per second, will be able to resolve the largest and most vigorous types of thunderstorm². But cloud physics on scales smaller than a grid cell will still have to be approximated, or

parametrized, using simplified equations.

Errors introduced by such parametrizations proliferate and infect calculations on larger scales. In climate simulators, these errors can be represented by introducing stochastic noise into the computational equations³. Hence, climate prediction is inherently probabilistic.

The main obstacle to building a commercially viable ‘exascale’ computer is not the flop rate itself but the ability to achieve this rate without excessive power consumption. Early estimates suggested that such computers would consume about 100 megawatts — the output of a small power station. A key challenge in recent years has been to make exascale computers more energy efficient.

Energy is required to perform basic arithmetic operations in computers. As microprocessors shrink to the nanometre scale, extra energy is needed to overcome thermal noise and even cosmic-ray strikes. By turning down the voltage, processors can be switched from deterministic to probabilistic calculators. For example, based on contemporary chip technology, a fourfold reduction in power can result in less than a 1% chance that a computational step will be incorrect⁴.

More importantly, energy is also used to move data from one part of a computer to another. The amount needed to accomplish this is proportional to the number of bits used to represent individual pieces of data. The ‘gold standard’ for variables taking real-number values is the 64-bit double-precision representation. Although supercomputers used in scientific computation also support 32-bit representations, they do not support representations with less than 32 bits, presumably because vendors perceive there to be little demand for such variable types.

There is no point, however, in being more deterministic or precise than is justified by the overall accuracy of the computational code; in the case of climate models, this accuracy is limited by errors from the parametrization schemes. Although 64-bit precision may be appropriate in representing variables associated with planetary-scale jet streams or weather systems that are hundreds of kilometres across, it is a waste of computing and energy resources to use this precision to represent smaller-scale circulations approaching the resolution limit of a climate model⁵. This is important because, collectively, the small-scale computations and sub-grid parametrization formulae dominate the cost of a climate simulation.

The energy liberated by not performing overly exact calculations could be put to more productive use. This requires a new type of supercomputer. Like current ones, such a machine would be massively

parallel, comprising many millions of individual processing units. A fraction of these would enable conventional, energy-intensive and deterministic high-precision computation. Unlike conventional computers, the remaining processors would be designed to take on low-energy probabilistic computation with lower-precision arithmetic, the degree of imprecision and inexactness being variable.

By using power more efficiently, a hybrid exascale supercomputer could extend the dynamic range of climate models to below the kilometre scale, allowing deep convective clouds to be well resolved. This would enable more reliable probabilistic predictions of Earth’s future climate.

WIDER BENEFITS

Inexact hybrid computing has the potential to aid modelling of any complex nonlinear multiscale system — from galactic and stellar evolution to tokamak plasmas and combustion in jet engines.

Living systems may already be aware of the benefits. The brain achieves its prodigious signal-processing capabilities using around 20 watts, less power than a typical light bulb. Axons — nerve fibres — and the ion channels that amplify the electrical signals that pass along them are of molecular dimensions and require little energy. The diameter of a typical human axon, 0.1 micrometres, is so small that the signals it contains are susceptible to thermal noise and hence to random fluctuations⁶. Although such noise is often considered to hinder the operation of the nervous system, in some circumstances it might offer an advantage⁷.

When considering computational systems that might mimic the brain, mathematician Alan Turing suggested that it would be “wise to include a random element in a learning machine” and provided a simple theoretical example to back up his claim⁸. It is now well known that adding noise can make algorithms more efficient⁹. For example, in the classic travelling salesperson problem — calculating the shortest possible route through multiple cities — adding a random noise component can reduce the overall computation time needed to find a solution⁹.

These considerations suggest that energy-efficient hybrid computing is indispensable for modelling the brain, and hence for understanding cognition. Indeed, it has been suggested that human creativity arises from a close synergy between low-energy computation (in which the operation of the brain is

susceptible to the brain’s own thermal noise) and higher-energy determinism (in which the implications of partially random cognitive jumps can be explored algorithmically in localized parts of the brain)¹⁰. If this is the case, then human creativity might be a by-product of evolutionary pressures to optimize the use of energy available to power the brain.

PARADIGM SHIFT

Supercomputer manufacturers are driven by commercial pressures. The type of computer architecture I envisage will be built only if there is sufficient demand from the scientific community.

Scientists who use the top tier of supercomputers across a range of disciplines need to assess the extent to which the conventional — energy-intensive — deterministic approach to computation is becoming a bottleneck. To do this, they need to quantify the impact of inherent uncertainties associated with approximations in their computational codes. They can then estimate how much the levels of computational inexactness and indeterminism can be increased before the impact of these factors exceeds that of the inherent uncertainties.

Computer vendors should begin by marketing processors and scientific computing libraries that make efficient use of mixed-precision representations of real-number variables. Joint funding from research councils and the private sector should be made available to catalyse these developments.

High-performance computation is rapidly overtaking traditional experimentation in many scientific disciplines. In designing the next generation of supercomputers, we must embrace inexactness if that allows a more efficient use of energy and thereby increases the accuracy and reliability of our simulations. ■

Tim Palmer is a Royal Society research professor of climate physics and co-director of the Oxford Martin Programme on Modelling and Predicting Climate at the University of Oxford, UK.
e-mail: tim.palmer@physics.ox.ac.uk

1. Mauritsen, T. & Stevens, B. *Nature Geosci.* **8**, 346–351 (2015).
2. Palmer, T. *Nature* **515**, 338–339 (2014).
3. Palmer, T. N. Q. J. R. *Meteorol. Soc.* **138**, 841–861 (2012).
4. Palem, K. V. *Phil. Trans. R. Soc. A* **372**, 20130281 (2014).
5. Düben, P. D. & Palmer, T. N. *Mon. Weath. Rev.* **142**, 3809–3829 (2014).
6. Faisal, A. A. & Laughlin, S. B. *PLoS Comput. Biol.* **3**, e79 (2007).
7. McDonnell, M. D. & Ward, L. M. *Nature Rev. Neurosci.* **12**, 415–426 (2011).
8. Turing, A. M. *Mind* **236**, 433–460 (1950).
9. Hoos, H. H. & Stützle, T. *Stochastic Local Search: Foundations and Applications* (Elsevier, 2005).
10. Palmer, T. N. & O’Shea, M. *Front. Comput. Neurosci.* **9**, 124 (2015).