Robust and Scalable Solution of Partial Differential Equations: Approaches from Mathematics and Computer Science

Peter Strazdins
Computer Systems Group,
Research School of Computer Science,
The Australian National University

collaborators: Markus Hegland, Richard Brent, Steve Roberts (MSI ANU), Alistair Rendell (RSCS ANU) ; Ross Nobes (Fujitsu Labs Europe)

Sandia National Laboratories, 11 Nov 2011

(slides available from http://cs.anu.edu.au/~Peter.Strazdins/seminars)

THE AUSTRALIAN NATIONAL UNIVERSITY

1 Talk Overview

- context: why we need new algorithms
- overview of the ANU Petascale Algorithms Project
- project drivers
- mathematical ideas: solution on multiple grids
- computational infrastructure
- applications: ANUGA and GENE
- performance analysis techniques
- conclusions and future work
- recent work in performance analysis
  - atmosphere simulation: the Met Office Unified Model
  - heart simulation: Chaste

2 Why We Need New Algorithms

- Kei supercomputer developed by Fujitsu and RIKEN has 80,000 processors
  - connected by a sophisticated network (TOFU)
- as synchronization at this scale is expensive, we need fundamentally asynchronous algorithms
- need a new approach relaxing computational dependencies and resilient against
  - missing data not available at the right time)
  - errors (due to very infrequent hard and software faults)

3 Overview of the ANU-FLE Petascale Algorithms Project

- investigate and develop new computational methods which underpin the deployment of emerging supercomputer hardware to the sciences
- co-developed with 2 applications in tsunami modelling & plasma physics
- aim to deliver scalability for very high levels of parallelism, asynchronous computation, delivering:
  - deliver scientific understanding of these new mathematical techniques
  - new techniques and software which demonstrates these techniques
  - training of higher research students in this crucial technology
- as well as the co-investigators, team to consist of 1 postdoc and 2–4 PhD students (under recruitment!)
- budget: ≈200K p.a. over 3 years from mid 2011
- broader collaboration with FLE researchers, the Open Petascale Libraries Project, and others . . .

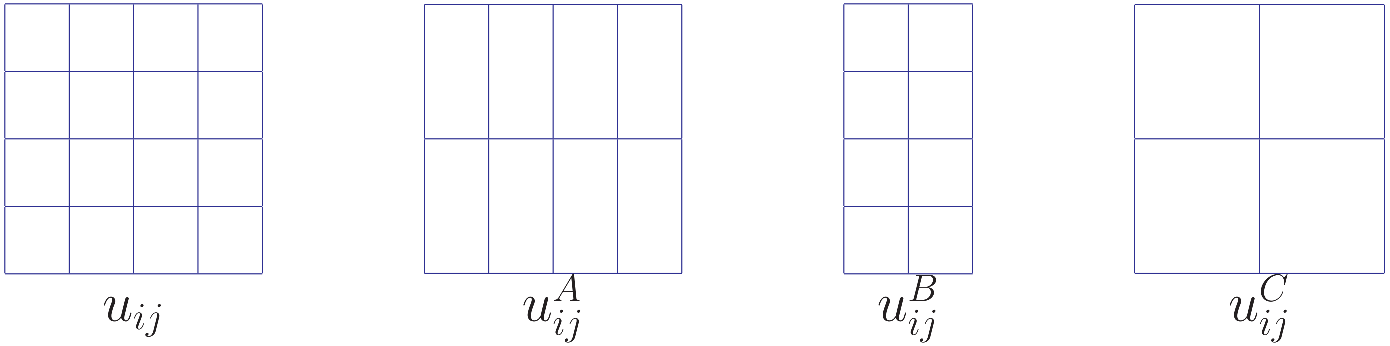
4 Project Drivers

- mathematics (first order hyperbolic PDEs of form $\frac{\partial u_i}{\partial t} = \nabla j_i(u) + s_i(u)$)
  - multidimensional approximation
  - wavelets
  - nonlinear approximation theory
- technology
  - large scale clusters
  - manycore cluster nodes
  - novel communication networks
- applications: transport of mass and momentum, turbulence, wide range of scales
  - tsunami (inundation) modelling - shallow water equations
  - plasma physics - Vlasov equations

Addressing such a challenge requires close collaboration between mathematicians, computer scientists and hardware manufacturers.

5 Robust techniques for time-dependent PDEs: Basic Ideas

- solve same problem on multiple different grids



$u_{11} \approx \gamma_A u_{11}^{A} + \gamma_B u_{11}^{B} + \gamma_C u_{11}^{C}$ general formula

$u_{11} \approx \frac{1}{2} u_{11}^{A} + \frac{1}{2} u_{11}^{B} - \frac{1}{4} u_{11}^{C}$ high accuracy

$u_{11} \approx \frac{1}{2} u_{11}^{B}$ if u_{11}^{A} not available

where $u_{11}^{[A,B,C]}$ is integral of u over top-left rectangle

- use wavelet analysis and nonlinear approximation theory (derive error bounds)
- develop new robust extrapolation techniques which are able to both correct errors and estimate missing data

6 Computational Infrastructure: Issues to be Addressed

- optimization of communication performance: optimal mappings of grids to nodes
- determining policies for the detection of faults – use wavelet theory to assess the severity of the faults
- resilient collective communications (not completely avoidable!)
- re-assignment of failed threads (each in charge of computing a sub-grid) to both new and existing nodes
- intra- and inter-node optimizations for specific kernels developed by the project
- load balancing techniques for large-scale multicore (destructive sharing)
- simulation of faults, under a realistic fault model
- models will be developed both for the energy and time costs of these computations over general multicore clusters

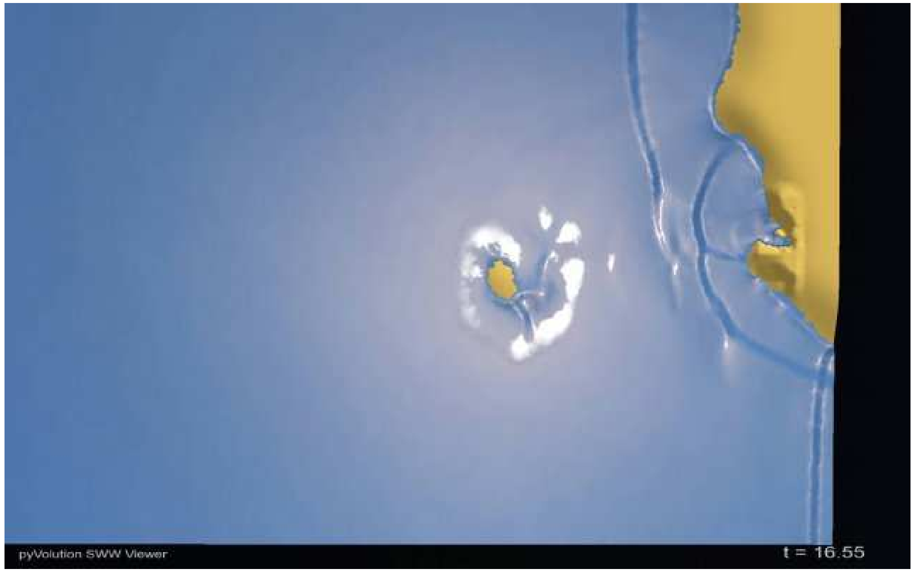
7 Computational Infrastructure: Kei-computer Issues

- network-specific optimisations for the TOFU network:
  - 6-D torus specific grid mappings
  - utilisation of some of the middle dimensions
  - explore hardware-supported collectives

  Note: this network has a lot of hardware fault tolerance!
- processor-specific optimisations for Sparc FX
  - programmable caches for kernel optimization
  - hardware-supported barriers
  - interaction with VISIMPACT low-level system software

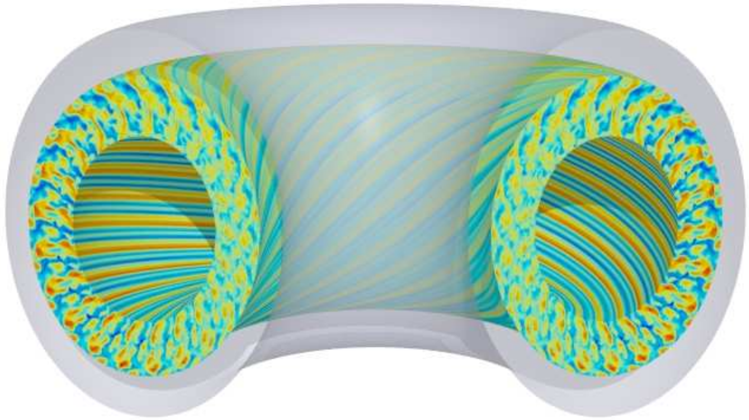
8 Applications - ANUGA

- ANUGA tsunami propagation and inundation modelling
  - collaboration with Geosciences Australia – Steve Roberts, Ole Nielsen
  - links with with OzAid and projects in the region
- shallow water wave equation, takes into account friction & bed elevation
  - triangles of variable size according to topography and interest
  - time step determined by triangle size and wave speed
  - compute fluxes across triangle edges using the Central Scheme
  - open source software: Python, C and MPI



THE AUSTRALIAN NATIONAL UNIVERSITY

9 Applications - GENE & Others



- GENE: Gyrokinetic Electromagnetic Numerical Experiment
  - plasma microturbulence code
  - multidimensional solver of Vlasov equation
  - fixed grid in five-dimensional phase space
- GENE: computes gyroradius-scale fluctuations and transport coefficients
  - hybrid MPI/OpenMP parallelization – high scalability to 2K cores
- other applications of potential interest:
  - atmosphere modelling:
    Met Office Unified Model (too complex for research?) WRF? CAM?
  - cardiac simulation: Chaste

10 Performance Evaluation

- important aspect of the project
- limited time allocation on some systems - needs to be 'efficient'
  - need to isolate a small representative section of the workload for detailed analysis
- need to determine:
  - profiles of communication: counts of MPI calls, attributable to major sections
  - patterns of stalls in communication and how these are handled by the infrastructure
  - memory hierarchy bottlenecks
  - data to build up performance models
- will existing tools like IPM be adequate?

11 Conclusions and Future Work

- new mathematical methods have a great promise to the area
  - redundancy can improve accuracy and/or permit resilience
- emerging petascale and beyond systems offer an exciting opportunity
- project is only just starting. . .
- really is a very considerable amount of work!
  - while ANUGA is relatively simple, still not trivial; others complex!
- need to collaborate widely / leverage efforts from elsewhere

Questions???