Running Large-Scale Ultrasound Simulations on Piz Daint with 512 Pascal GPUs

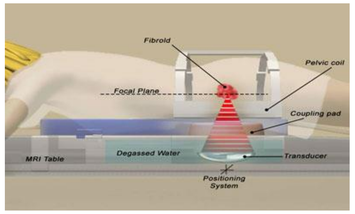
Filip Vaverka[1], Bradley E. Treeby[2], Jiri Jaros[1]

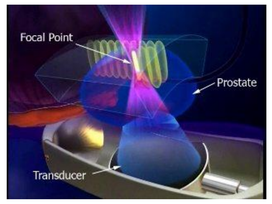
[1]Faculty of Information Technology, Brno University of Technology, CZ
[2]Department of Medical Physics and Bioengineering, University College London, UK

k-Wave
A MATLAB toolbox for the time-domain simulation of acoustic wave fields

1 Overview

High-intensity focused ultrasound (HIFU) is an emerging non-invasive cancer therapy that uses tightly focused ultrasound waves to destroy tissue cells through localised heating. The treatment planning goal is to select the best transducer position and transmit parameters to accurately target the tumour. The path of the ultrasound waves can be predicted by solving acoustic equations based on mass, momentum and energy conservation. However, this is a computationally difficult problem because the domain size is very large compared to the acoustic wavelength.





2 Nonlinear Ultrasound Wave Propagation in Tissue

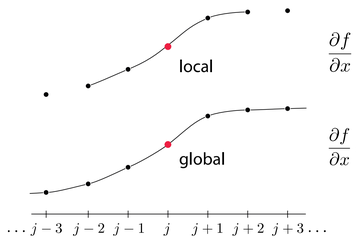
The governing equations must account for the nonlinear propagation of ultrasound waves in tissue, which is a heterogeneous and absorbing medium. Accurate acoustic absorption is critical for predicting ultrasound dose under different conditions.
The required acoustic equations can be written as:

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho_0}\nabla p \qquad \textit{momentum conservation}$$

$$\frac{\partial \rho}{\partial t} = -\left(2\rho + \rho_0\right)\nabla \cdot \mathbf{u} - \mathbf{u} \cdot \nabla \rho_0 \qquad \textit{mass conservation}$$

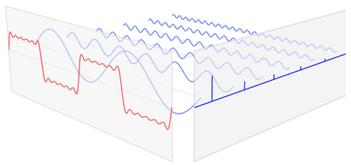
$$p = c_0^2\left(\rho + \mathbf{d} \cdot \nabla \rho_0 + \frac{B}{2A}\frac{\rho^2}{\rho_0} - \Pi\rho\right) \qquad \textit{pressure-density relation}$$

These equations are discretized using the *k*-space pseudo-spectral method and solved iteratively. This reduces the number of required grid points per wavelength by an order of magnitude compared to finite element or finite difference methods. For uniform Cartesian grids, the gradients can be calculated using the fast Fourier transform (FFT).



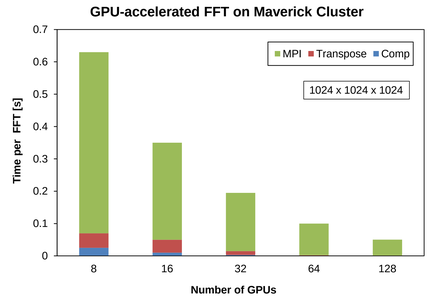
$$\frac{\partial f}{\partial x} \approx \frac{\frac{1}{12}f^{j-2} - \frac{2}{3}f^{j-1} + \frac{2}{3}f^{j+1} - \frac{1}{12}f^{j+2}}{\Delta x}$$

$$\frac{\partial f}{\partial x} \approx \mathbb{F}^{-1}\left\{ik_x \mathbb{F}\left\{f\right\}\right\}$$



3 3D Fourier Transforms over Multiple GPUs

Running 3D FFTs across multiple GPUs overloads the interconnection network by performing distributed matrix transpositions. The efficiency of distributed spectral methods can then drop below 1%.

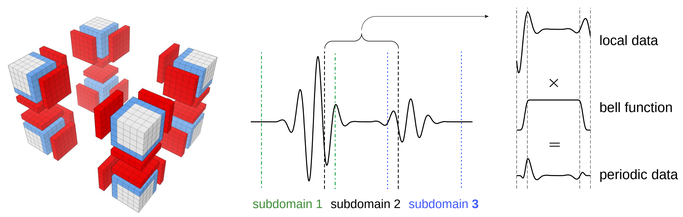


4 Local Fourier Basis Decomposition

Local domain decomposition reduces the communication burden by partitioning the domain into a grid of local subdomains where gradients are calculated locally and the global communication is replaced by the nearest-neighbor overlap exchange.

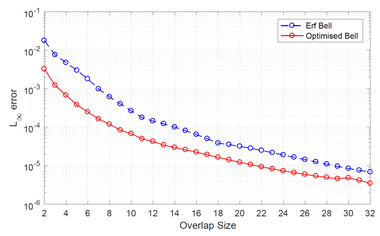
The gradient calculation with the overlap on an *i*-th subdomain reads as follows (*b* is a bell function smoothening the subdomain interface):

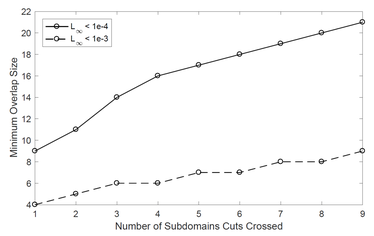
$$\frac{\partial p_i}{\partial t} = \mathbb{F}^{-1}\{ik_i\mathbb{F}(b \cdot p_i)\}$$



5 Local Fourier Basis Accuracy

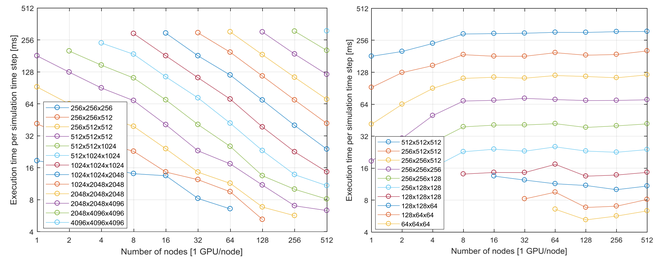
Since the gradient is not calculated on the whole data, numeric error is introduced. Its level can be controlled by the shape of the bell function and thickness of the overlap region. An overlap size of 8 grid points is sufficient to maintain the L_∞ error below 0.1% even after the wave has crossed 8 subdomain boundaries. For 3D decompositions, this corresponds to $9^3 = 729$ local subdomains.





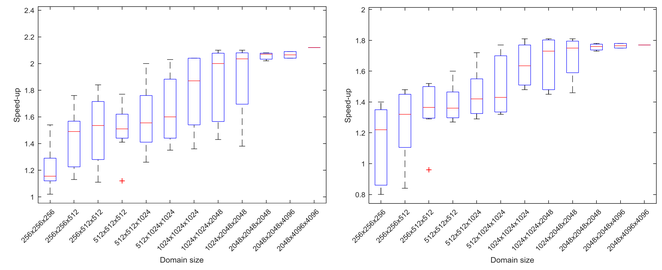
6 Performance Investigation

Strong and weak scaling of ultrasound simulations with local domain decomposition and overlap size of 8 on Piz Daint with 512 GPUs and aries network. Domain sizes between 256^3 and 4096^3 grid points.



7 Overlap Size and Performance Trade-off

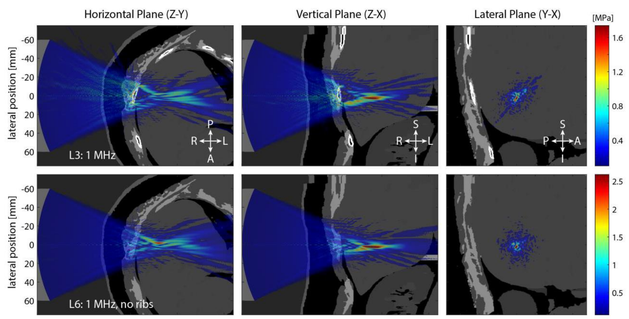
Box plots of scaling curves gradients (speed-up when doubling the number of subdomains) for overlap sizes of 4 and 16 on Piz Daint.



8 Conclusions

By numerical optimization of the bell function and careful overlap size selection, we achieved efficiency between 90 and 100% for 512 GPUs of Piz Daint.

When compared with the global slab decomposition on a Haswell based cluster, the GPU solution offers more than 100x higher performance and 12x lower simulation cost. This enables us to run realistic HIFU treatment plans in terms of hours.



This work was supported by The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project IT4Innovations excellence in science - LQ1602 and by the IT4Innovations infrastructure which is supported from the Large Infrastructures for Research, Experimental Development and Innovations project IT4Innovations National Supercomputing Center - LM2015070. This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID pr14. This work was also supported by the Engineering and Physical Sciences Research Council, UK, grant numbers EP/L020262/1 and EP/M011119/1. This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020 ICT 2016-2017 under grant agreement No 732411 and is an initiative of the Photonics Public Private Partnership.