

New research provides effective battle planning for supercomputer war

November 11 2010



This is professor Stephen Jarvis of the University of Warwick. Credit: University of Warwick

New research from the University of Warwick, to be presented at the World's largest supercomputing conference next week, pits China's new No. 1 supercomputer against alternative US designs. The work provides crucial new analysis that will benefit the battle plans of both sides, in an escalating war between two competing technologies.

Professor Stephen Jarvis, Royal Society Industry Fellow at the

University of Warwick's Department of Computer Science, will tell some of the 15,000 delegates in New Orleans next week, how general-purpose GPU (GPGPU) designs used in China's 2.5 Petaflops Tianhe-1A fare against alternative [supercomputing](#) designs employed in the US; these use relatively simpler processing cores brought together in parallel by highly-effective and scalable [interconnects](#), as seen in the IBM BlueGene architectures.

Professor Jarvis says that:

"The 'Should I buy GPGPUs or BlueGene' debate ticks all the boxes for a good fight. No one is quite sure of the design that is going to get us to Exascale computing, the next milestone in 21st-century computing, one quintillion floating-point operations per second (10^{18}). It's not simply an architectural decision either – you could run a small town on the power required to run one of these supercomputers and even if you plump for a design and power the thing up, programming it is currently impossible."

Professor Jarvis' research uses mathematical models, benchmarking and simulation to determine the likely performance of these future computing designs at scale:

"At Supercomputing in New Orleans we directly compare GPGPU designs with that of the BlueGene. If you are investing billions of Dollars or Yuan in supercomputing programmes, then it is worth standing back and calculating what designs might realistically get you to Exascale, and once you have that design, mitigating for the known risks – power, resilience and programmability."

Professor Jarvis' paper uses mathematical modeling to highlight some of the biggest challenges in the supercomputing war. The first of these is a massive programming/engineering gap, where even the best computer programmers are struggling to use even a small fraction of the

computing power that the latest supercomputing designs have and, which will continue to be a problem without significant innovation. Professor Jarvis says:

"if your application fits, then GPGPU solutions will outgun BlueGene designs on peak performance" – but he also illustrates potential pitfalls in this approach – "the Tianhe-1A has a theoretical peak performance of 4.7 Petaflops, yet our best programming code-based measures can only deliver 2.5 Petaflops of that peak, that's a lot of unused computer that you are powering. Contrast this with the Dawn BlueGene/P at Lawrence Livermore National Laboratory in the US, it's a small machine at 0.5 Petaflops peak [performance], but it delivers 0.415 Petaflops of that peak. In many ways this is not surprising, as our current programming models are designed around CPUs."

But the story doesn't end there. "The BlueGene design is not without its own problems. In our paper we show that BlueGenes can require many more processing elements than a GPU-based system to do the same work. Many of our scientific algorithms – the recipes for doing the calculations – just do not scale to this degree, so unless we invest in this area we are just going to end up with fantastic machines that we can not use."

Another key problem identified by the University of Warwick research is the fact that in the rush to use excitingly powerful GPGPUs, researchers have not yet put sufficient energy into devising the best technologies to actually link them together in parallel at massive scales.

Professor Jarvis' modeling found that small GPU-based systems solved problems between 3 and 7 times faster than traditional CPU-based designs. However he also found that as you increased the number of processing elements linked together, the performance of the GPU-based systems improved at a much slower rate than the BlueGene-style

machines.

Professor Jarvis concludes that:

"Given the crossroads at which supercomputing stands, and the national pride at stake in achieving Exascale, this design battle will continue to be hotly contested. It will also need the best modeling techniques that the community can provide to discern good design from bad."

More information: A PDF of the paper can be found at:
www2.warwick.ac.uk/fac/sci/dcs...tions/pubs/sc-lu.pdf

Provided by University of Warwick

Citation: New research provides effective battle planning for supercomputer war (2010, November 11) retrieved 26 April 2024 from <https://phys.org/news/2010-11-effective-supercomputer-war.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.