



Environmental Considerations when Constructing GPGPU Codes

Issues for Performance and Benchmarking

Daniel Goodman
Daniel.Goodman@man.ac.uk



Codes Used

- Horn-Shunk algorithm for motion approximation
- Simulator and beam former for the Square Kilometer Array (SKA)
- Astronomical Data Analysis Algorithms
- Matrix Matrix Multiplication

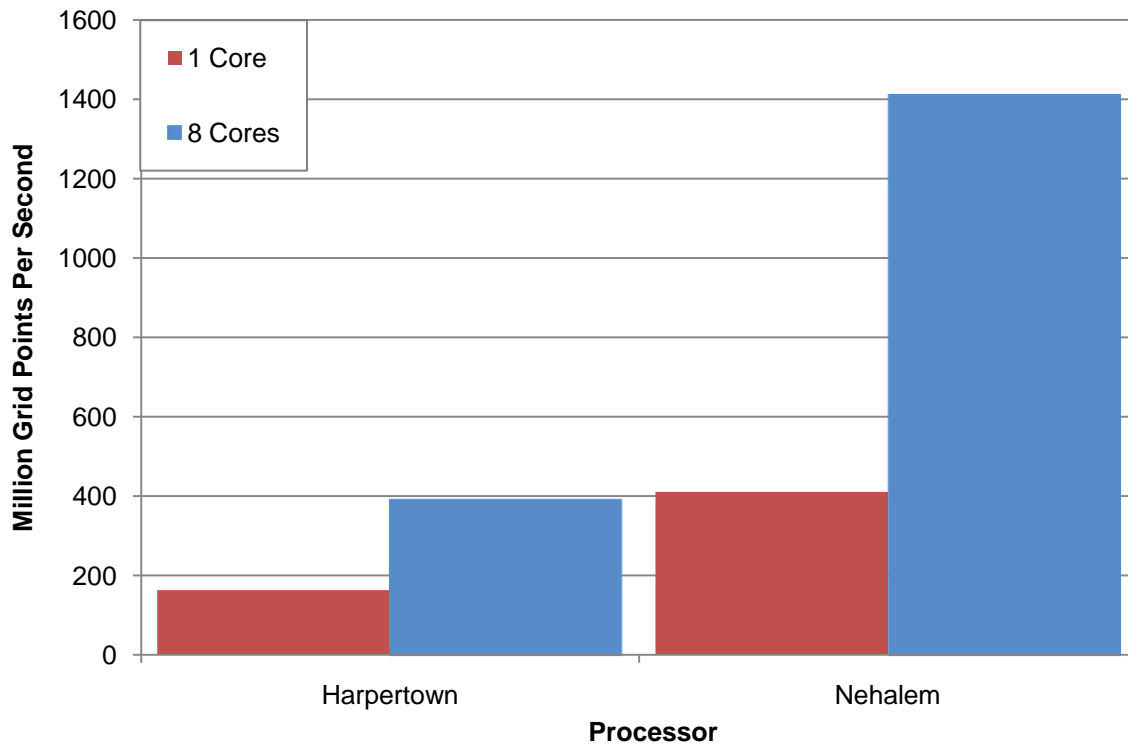


Mixed Mode Computation

- To take full advantage of GPU systems, both the CPU and GPU must be used at the same time
- Experimental SKA code used the CPU to construct a matrix for analysis and the GPU to perform the analysis
- Single Precision arithmetic can result in 10X slowdown for CPU code



Effect of Different CPU's



Number of astronomical data points gridded per second

- The improvement across multiple cores is much better than the improvement for a single core
- GTX260 achieves almost 1.6 billion

Results courtesy of Tim Cornwell

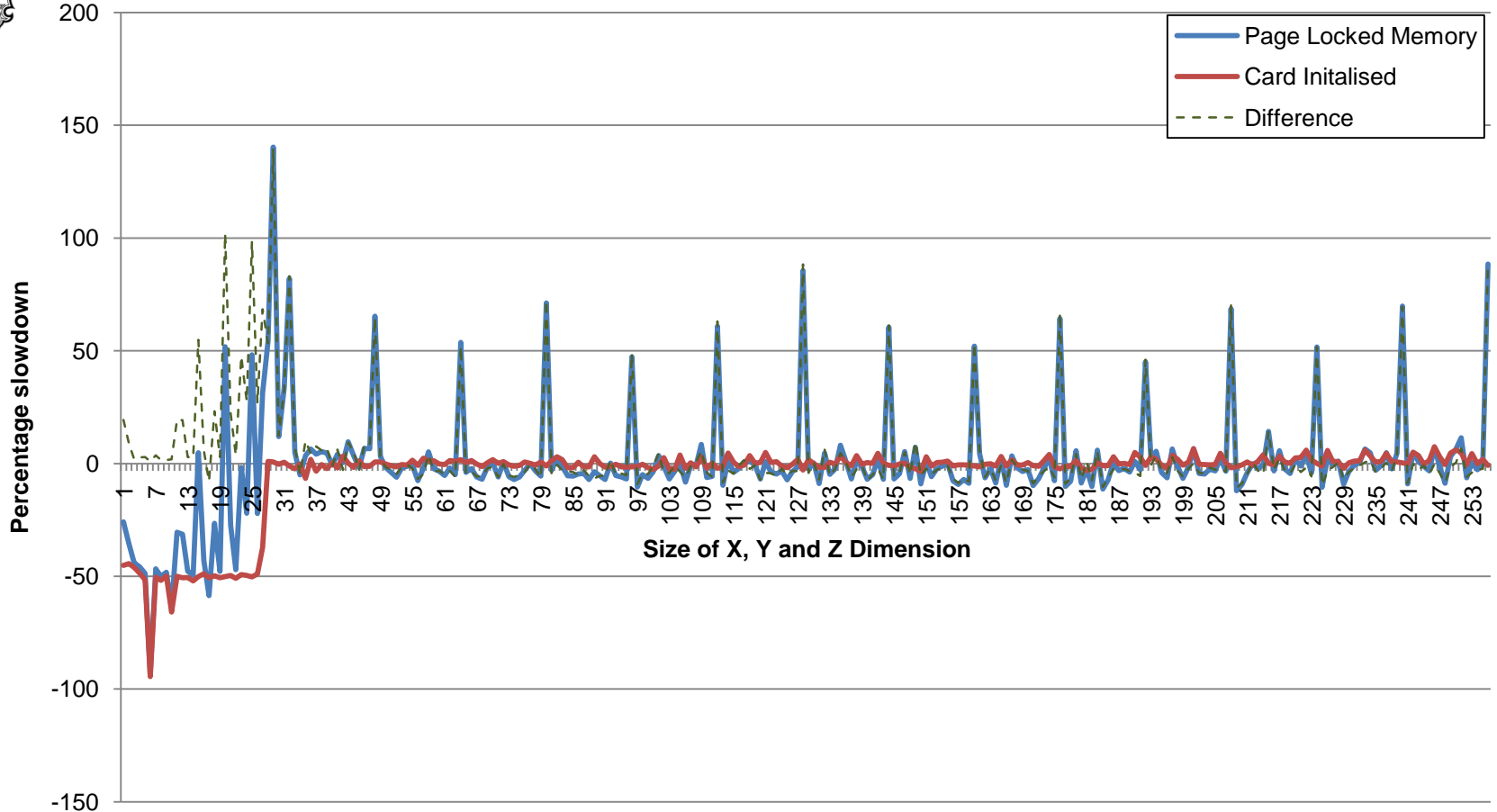


Effect of GPU Context

- Cuda code creates a context
- This is used by:
 - Card Initialisations
 - Page Locked Memory
 - Kernel Executions
- This context effects the CPU code as well

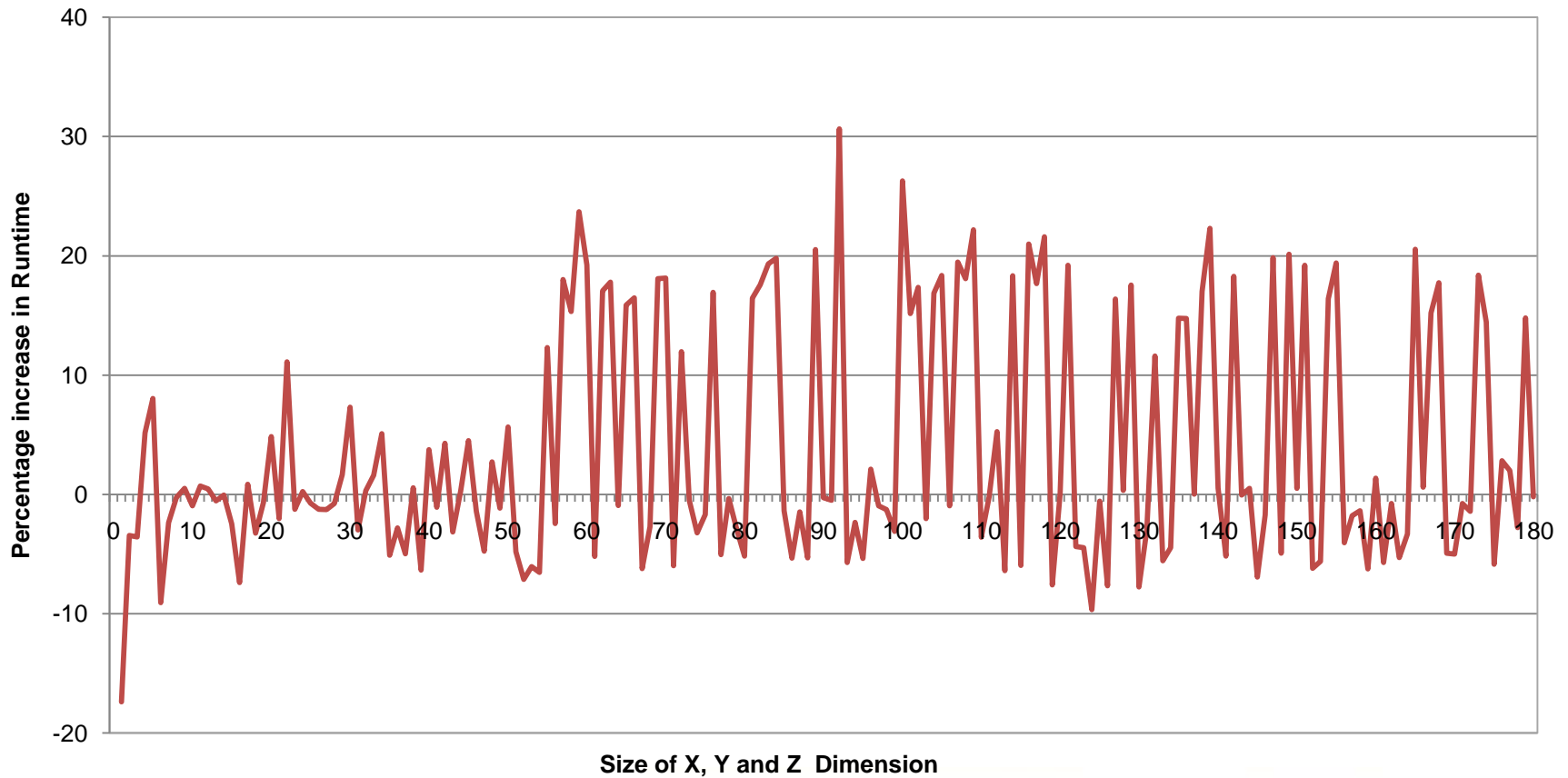


Effect of GPU Context





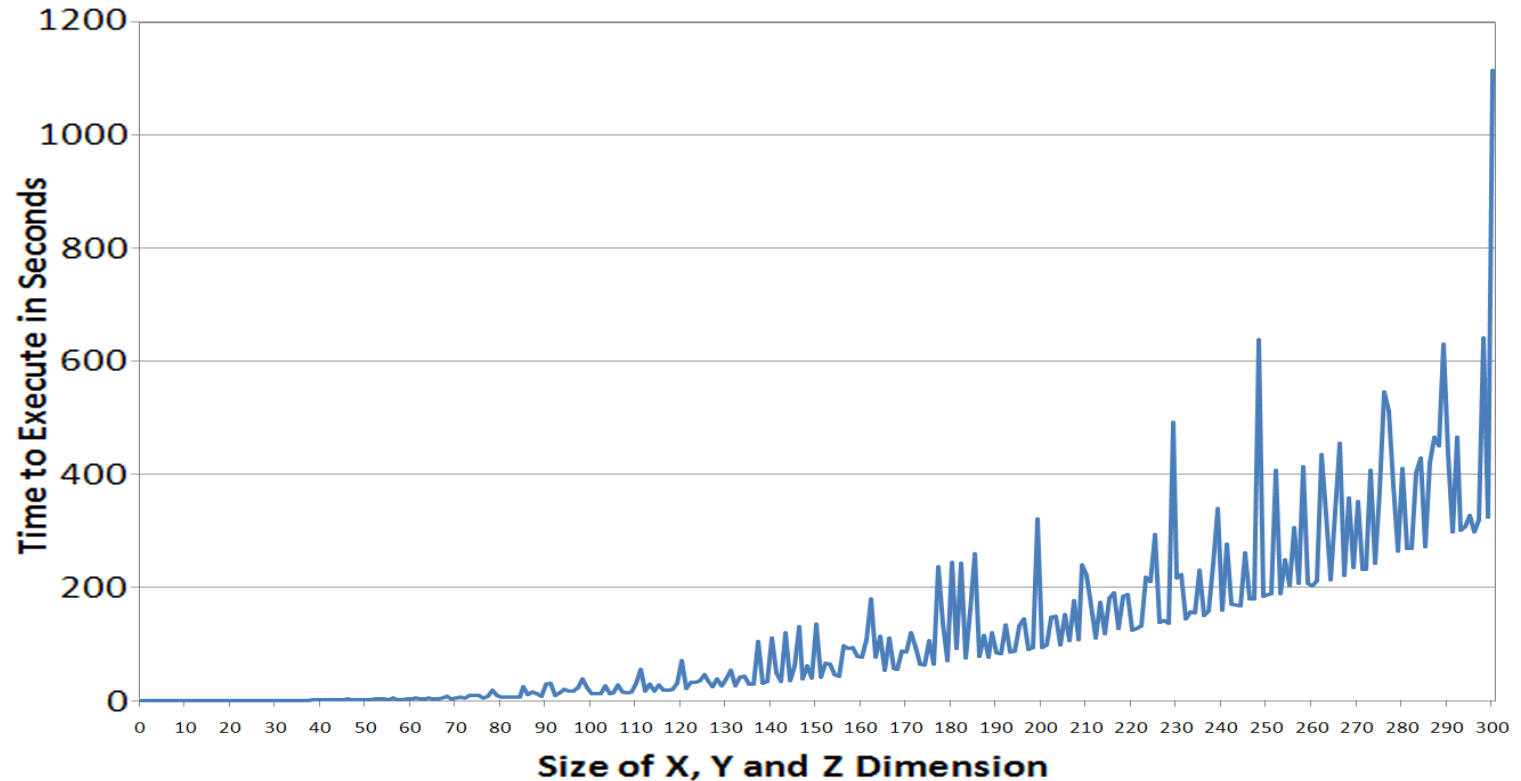
Effect of Kernel Execution





Bad Sizes Go Both Ways

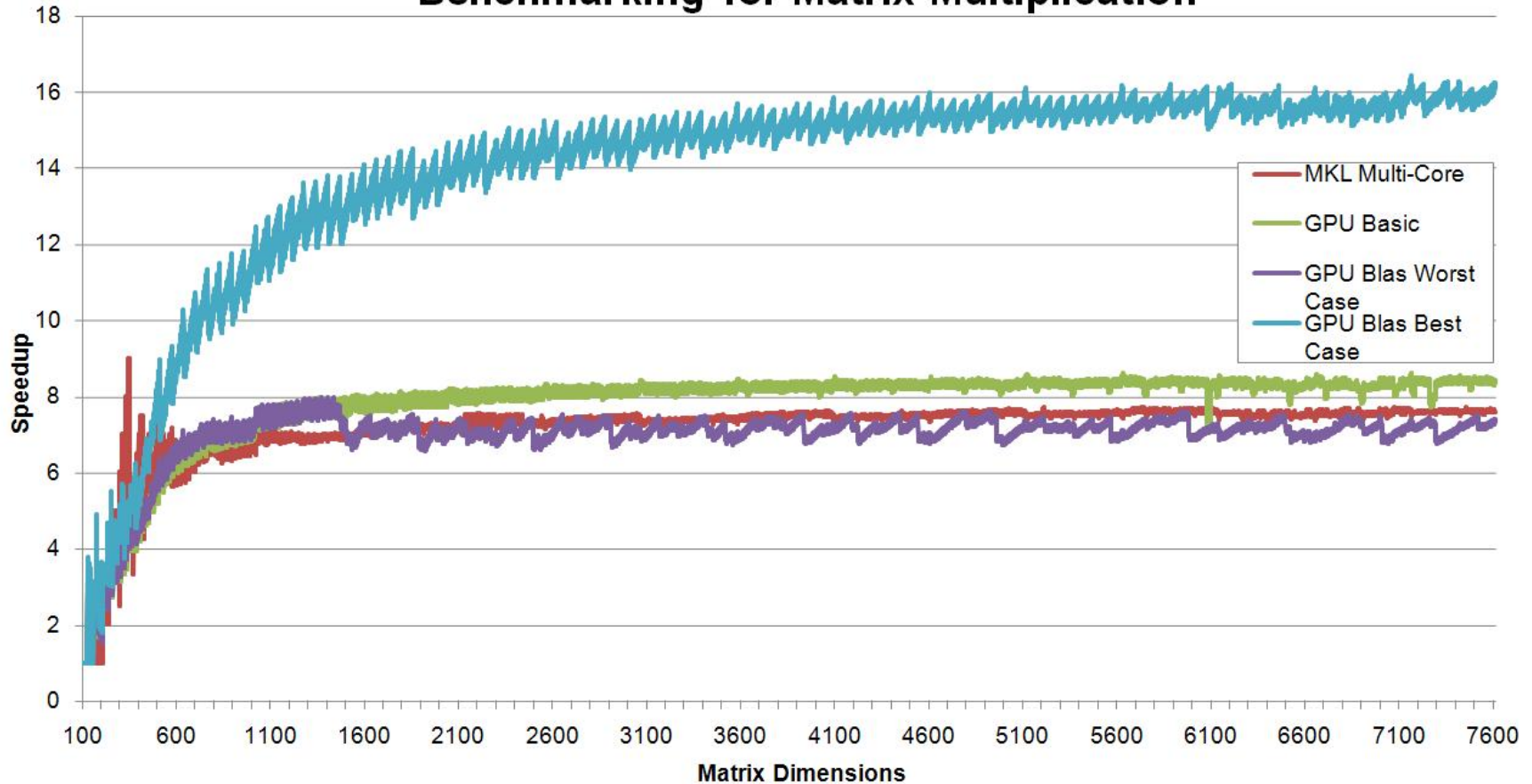
Time to perform computation on CPU





Effect of Data Size

Benchmarking for Matrix Multiplication





Conclusions

- To get best performance it will be necessary to use both the CPU and the GPU
- The relationship between the two is more complex than you might expect with interactions in unseen locations
- Careful thought needs to be given to ensure the best possible interaction in this moving environment
- Much of this may have changed as GPU's progress