



ENERGY EFFICIENCY IN HIGH PERFORMANCE HETEROGENEOUS COMPUTING

Brian Bergner & Dr. Rong Ge



INTRODUCTION

- High performance computing (HPC) refers to the complicated computer systems used in fields such as molecular modeling and weather forecasting
- Computing systems today need to be more energy efficient and faster than existing systems
- As Moore's Law predicted transistor densities are still doubling every 1.5 years
- Clock speeds however are limited by power availability, heat dissipation, and current leakage
- Applications can be optimized to use multi core hardware and coprocessor based accelerators
- Large software applications specifically those requiring repeated matrix operations benefit from this optimization
- Parallelism and hardware efficiency is the future of performance gains through the use of multi-core cluster on a chip technology
- Optimal points in computer performance can give you a competitive advantage with the best possible time-to-solution and minimizes waste of resources

PROBLEM & OBJECTIVE

- Many of the world's supercomputers, terascale and petascale systems, are designed to perform as many Floating Point Operations Per Second (FLOPS) as possible. Performance improvements in each node can have ripple effects across an entire system.
- The heterogeneous computing node studied includes an NVIDIA Tesla GPU (Graphic Processing Units) and an Intel Xeon Phi MIC (Many Integrated Core) coprocessor as system accelerators.
- Our goal is to systematically find optimal load balancing points given specific resource constraints that maximize energy efficiency while meeting requirements in throughput. The proposed approach is through a system of performance tests and subsequent analysis of energy sensitivity and power consumption. These results can be used in evaluating current and future computing hardware and interfaces to identify appropriate combinations.
- A heterogeneous system containing a CPU, a GPU and an Intel MIC with a PCI Express interface is being studied based on a BLAS (Basic Linear Algebra Subprograms) operation.
- Results indicate that a heterogeneous computing system is able to provide more energy-efficient solutions to scientific computing with various performance demands as the improvement of system energy efficiency is sensitive to the system components.

HARDWARE

- 'Ivy' is an Intel Based Server with Dual Ivy Bridge Intel Xeon E5-2670 v2 CPUs
- Each CPU has 10 Cores clocked at 2.56GHz with a 25MB Intel Smart Cache
- The NVIDIA GPU accelerator card is connected via the PCI-Express bus is a Tesla K20c which features 2688 CUDA Cores
- The Intel Xeon Phi Coprocessor is connected via the PCI-Express bus and has 7120D has 61 Cores clocked at 1.238GHz and sharing a 30.5MB L2 Cache
- External power meter records the total energy being supplied to the system

IMPLEMENTATION & TESTING

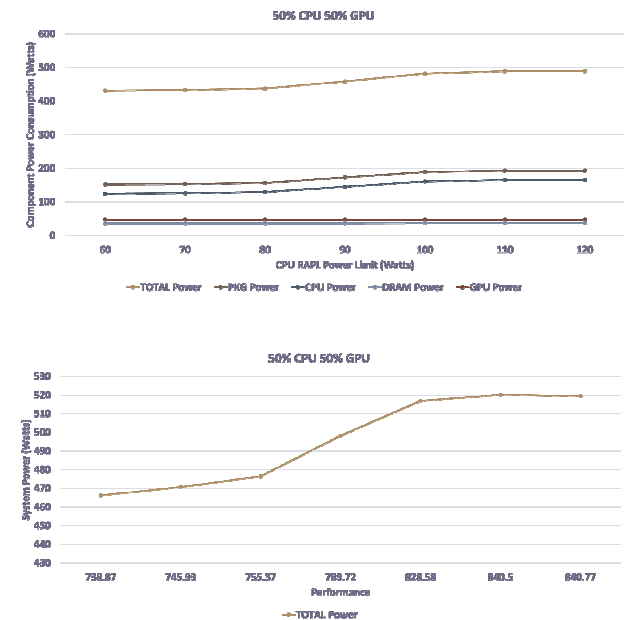
- Jobs are submitted to the 'Ivy' system for processing from a remote computer
- Power measurements are sampled at regular intervals and recorded while the Matrix operations are processing on the CPU and Accelerator
- Power averages are calculated and recorded along with performance metrics based on time of job submission and completion
- Each test result creates a data point for further analysis
- For each test only one variable is modified at a time
- Variables tested Include:
 - Percent of Matrix Computation per Component
 - Software Power Limits Applied to CPU Package
- Further variables can include:
 - Different Accelerator Hardware
 - Matrix Operation Size
 - Matrix Operation Type
 - # of Iterations of Matrix Operation
 - Software Based Power Limits Applied to Accelerator
 - Software Power Limits Applied to CPU Core
 - Software Power Limits Applied to DRAM
 - Hardware Based Power Limits

| CPU Percent | GPU Percent | Matrix Size | Itera | Perf | Power | CPU Performance | GPU Performance | Total Time | CPU Time | GPU Time | CPU Speed | GPU Speed | TOTAL Power | CPU Power | GPU Power | DRAM Power | SFU Power |
|-------------|-------------|-------------|-------|------|--------|-----------------|-----------------|------------|----------|----------|-----------|-----------|-------------|-----------|-----------|------------|-----------|
| 100 | 0 | 128000 | 10 | 80 | 348.89 | 348.89 | 0.00 | 123.76 | 123.76 | 0.00 | 1000000 | 708.00 | 484.74 | 148.84 | 24.76 | 88.89 | 47.26 |
| 100 | 0 | 128000 | 10 | 70 | 368.89 | 368.89 | 0.00 | 123.40 | 123.40 | 0.00 | 1000000 | 708.00 | 484.82 | 148.89 | 23.40 | 88.89 | 47.89 |
| 100 | 0 | 128000 | 10 | 60 | 371.35 | 371.35 | 0.00 | 123.88 | 123.88 | 0.00 | 1000000 | 708.00 | 487.04 | 146.67 | 139.27 | 88.89 | 47.85 |
| 100 | 0 | 128000 | 10 | 50 | 388.24 | 388.24 | 0.00 | 107.61 | 107.61 | 0.00 | 1000000 | 708.00 | 488.87 | 139.89 | 148.89 | 88.89 | 47.85 |
| 100 | 0 | 128000 | 10 | 40 | 423.09 | 423.09 | 0.00 | 101.79 | 101.79 | 0.00 | 1000000 | 708.00 | 488.89 | 148.89 | 148.89 | 87.04 | 47.85 |
| 100 | 0 | 128000 | 10 | 30 | 451.27 | 451.27 | 0.00 | 98.54 | 98.54 | 0.00 | 1000000 | 708.00 | 488.89 | 148.89 | 148.89 | 87.11 | 47.85 |
| 100 | 0 | 128000 | 10 | 20 | 423.48 | 423.48 | 0.00 | 98.38 | 98.38 | 0.00 | 1000000 | 708.00 | 488.74 | 148.82 | 148.82 | 87.04 | 47.85 |
| 80 | 20 | 128000 | 10 | 80 | 368.89 | 788.87 | 1090.47 | 94.78 | 94.78 | 13.89 | 1000000 | 708.00 | 491.28 | 148.89 | 138.89 | 87.78 | 99.89 |
| 80 | 20 | 128000 | 10 | 70 | 371.37 | 748.89 | 1080.47 | 88.38 | 88.38 | 13.89 | 1000000 | 708.00 | 478.84 | 147.89 | 128.89 | 84.74 | 91.89 |
| 80 | 20 | 128000 | 10 | 60 | 371.37 | 788.87 | 1080.47 | 88.38 | 88.38 | 13.89 | 1000000 | 708.00 | 478.84 | 148.89 | 138.89 | 84.84 | 91.89 |
| 80 | 20 | 128000 | 10 | 50 | 388.24 | 788.87 | 1090.47 | 84.11 | 84.11 | 13.89 | 1000000 | 708.00 | 491.24 | 148.89 | 138.89 | 83.89 | 98.89 |
| 80 | 20 | 128000 | 10 | 40 | 423.09 | 888.89 | 1090.47 | 88.89 | 88.89 | 13.89 | 1000000 | 708.00 | 501.89 | 177.89 | 148.89 | 84.78 | 98.49 |
| 80 | 20 | 128000 | 10 | 30 | 423.09 | 848.89 | 1080.47 | 88.89 | 88.89 | 13.89 | 1000000 | 708.00 | 501.89 | 177.89 | 148.89 | 84.89 | 98.49 |
| 80 | 20 | 128000 | 10 | 20 | 423.09 | 848.89 | 1080.47 | 88.89 | 88.89 | 13.89 | 1000000 | 708.00 | 501.89 | 177.89 | 148.89 | 84.89 | 98.49 |
| 60 | 40 | 128000 | 10 | 80 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 70 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 60 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 50 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 40 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 30 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |
| 60 | 40 | 128000 | 10 | 20 | 368.89 | 1090.47 | 0.00 | 48.89 | 48.89 | 0.00 | 48.89 | 1000000 | 708.00 | 461.28 | 48.89 | 24.76 | 138.89 |

RESULTS & ANALYSIS

Results can be analyzed by asking:

- How much performance do we gain by parallelizing the code?
- Are there ways to gain energy efficiency by offloading only certain functions to the accelerator?
- Do specific BLAS operations perform better on specific accelerator hardware?
- How does Matrix size effect specific hardware efficiency?



SOFTWARE & CONFIGURATION

- 'Ivy' runs a Linux variant (CentOS 6.5) and is configured to measure and limit CPU power usage using Intel's Running Average Power Limit (RAPL) driver
- Intel Manycore Platform Software Stack (MPSS) 3.2 is used to offload computation to the Intel Xeon Phi Coprocessor
- Tests are run using matrix multiplication operations built using Basic Linear Algebra Subsystems (BLAS), which is the generally accepted building block standard for Linear Algebra functions
- cuBLAS is used as a library for GPU accelerated and offloaded subroutines

ACKNOWLEDGEMENTS

- Rong Ge, Ryan Vogt, Jahangir Majumder, Arif Alam, Martin Burtcher and Ziliang Zong. *Effects of Dynamic Voltage and Frequency Scaling on a K20 GPU*, Proceed
- Rong Ge, Xizhou Feng, Shuaiwen Song, Hung-Ching Chang, Dong Li, Kirk W. Cameron, PowerPack: Energy profiling and analysis of High-Performance Systems and Applications, IEEE Transactions on Parallel and Distributed Systems, Vol 21, No.5, 658-671 (2010). Proceedings of the 42nd International Conference on Parallel Processing. 2013
- Jeffers, Jim, and James Reinders. *Intel Xeon Phi Coprocessor High-performance Programming*.