Advanced compilation techniques, architectures and algorithms for Nuclear Physics calculations

Frank T. Winter, Bálint Jó, Robert G. Edwards, David G. Richards (presenter), Jefferson Lab
fwinter@jlab.org, bjoo@jlab.org, edwards@jlab.org, dgr@jlab.org

Introduction: QDP-JIT/PTX

The current diversity in computing architectures makes Domain Specific Language (DSL) approaches to constructing codes highly attractive. Users can write code using DS constructs, and have the DSL implementation ensure the best mapping onto the available system. The success of this approach is demonstrated by the QDP-JIT/PTX implementation of the QDP++ domain specific framework for Lattice QCD. The Chroma application, built over the QDP-JIT/PTX implementation is successfully accelerated on NVIDIA GPU based systems. In combination with the QUDA optimized solver library enabled Chroma on large scale GPU based systems. This, in combination with the QUDA optimized solver library enabled Chroma on large scale GPU based systems such as Titan at OLCF and Blue Waters at NCSA.

The Future: QDP-JIT/LLVM

Rearchitecting QDP-JIT, to generate Internal Representation (IR) for the LLVM compiler framework allows the JIT approach to be harnessed also for non-GPU targets. Generating LLVM IR has several advantages: efficient code generation for several targets (x86, PowerPC, GPU), as well as advanced code transformations and optimizations (fusion, fission, tiling, etc.)

Multigrid Solvers

Correlation functions for Hadron Spectroscopy calculations are computed from quark line diagrams. Each quark line is a quark propagator and is the result of solving the QCD Dirac Equation. Each configuration requires O(1M) solutions of the Dirac Equation. This motivates the use of optimized solver algorithms, such as recently developed Algebraic Multi-Grid Solvers

Xeon Phi and x86 Optimization

While the JIT approach is successful, there is always room for highly optimized, hand tuned solvers. We have continued our work on optimized solvers for Xeon Phi and Xeon (Joo et al. ISC’13). Our latest code features double precision and also 16 bit up-downconversion on Xeon Phi. Communication can now be performed in all 4-dimensions.

Conclusion

We have shown progress both on applying advanced compilation techniques to our code and exploiting advanced algorithms and architectures. These efforts enable our calculations on current systems and are paving the way for the next generation of extreme scale platforms.