Werner Benger,
Andreas Gerndt,
Simon Su,
Wolfram Schoor,
Michael Koppitz,
Wolfgang Kapferer,
Hans-Peter Bischof,
and Massimo Di Pierro

Proceedings of the 6th High-End Visualization Workshop

Open issues in visualization
with special concentration on applications in
astrophysics, numerical relativity, computational fluid dynamics
and high-performance computing

December 8th- 12th, 2010
Obergurgl, Tyrol, Austria

http://vizworkshop.cct.lsu.edu/viz2010/

Article 2

Visualization Workflow for Lattice QCD

Brian Schinazi, Yaoqian Zhong, Massimo Di Pierro
School of Computing, College of Computing and Digital Media, DePaul University
243 S Wabash Avenue, Chicago, IL, USA

Vis is a web based application that implements Software as a Service for Lattice QCD computations. Lattice QCD is a numerical approach to the mathematical model that describes quarks and gluons, the consituents of protons, neutrons and many other composite particles. Lattice QCD computations are implemented as Markov Chain Monte Carlo. Vis allows to store this data, explore it, schedule computing jobs using a local or remote PBS cluster, and schedule visualization jobs using VisIt as back-end. All the major operations of Vis can be performed via the web-based interface as well as scripted. Vis provides an access control mechanism and strong security features. It is a single platform that may allow physicists to collaborate better by sharing their data online.

2.1 Introduction

Lattice QCD [Massimo Di Pierro, 2006] is a numerical approach to the study of quarks, the elementary constituents of protons, neutrons and other forms of matter. In 1968, the study of physics reached a turning point when the structure and interactions of all known particles were described by a single mathematical expression, known as the Standard Model Lagrangian [Novaes, 1999]. Since that time, predictions based on the Standard Model have been extremely accurate, and have been able to account for the results of every high energy physics experiment.

Still, physicists continue to explore nature at smaller and smaller scales and to look for a breakdown of the model, manifested as a discrepancy between predictions and experiments. This would be a major discovery.

The part of the Standard Model that describes quarks specifically is called Quantum Chromodynamics [Altarelli, 2002]. This constitutes perhaps the most fascinating part of the Standard Model, as quarks are the only elementary particles subject to the strong nuclear force – a highly non-linear interaction that allow quarks to bind together in complex structures. Practically all of the composite particles we see in experiments are made up of quarks.

The goal of Lattice QCD is twofold: to compute from first principles the properties (such as masses and lifetimes) of these composite particles, and to extract fundamental parameters of QCD (such as particles' masses) from a comparison of theory with experiment.

Typical computations consist of taking a small portion of space ($10^{-15}m$ of side) and its evolution over a short period of time ($10^{-23}s$), and then performing a Markov Chain Monte Carlo (MCMC) simulation of all of its possible evolutions (1000 histories). We call the data saved at each MCMC step a gauge configuration. Next, correlation functions are measured over all simulated evolutions of the system. It can be proven that such an algorithm is equivalent to simulating a quantum-mechanical system. Finally, observable quantities are extracted from the correlation functions.

Until recently, visualization techniques have not been used in the study of QCD. The main reason is that the objects being computed have no obvious correspondence with physical 3D objects. The content of the portion of space which is simulated contains purely random data, since each data set is just a step of a MCMC. The physics is encoded in the probability distribution used to generate the MCMC, and not in the data itself.

We believe that there are some useful applications of visualization techniques to Lattice QCD: they can be used for didactic purposes, they can be used to better understand the behavior of the MCMC algorithms, and they can be used to detect certain types of error in the computation.

One type of error we are interested in is a systematic one: the possible long auto-correlation of topological charge distribution. The portion of space-time that is simulated contains a field that can be thought of as the chromo-electro-magnetic field of gluons, the particles that mediate interactions between quarks (analogous to the electro-magnetic field being represented by photons, but in this case having three types of charges). One can also associate a local topological charge density to the field. It should be noted that these fields all live in 4D and therefore must be projected on to 3D in order to be visualized.

The elementary steps of Lattice QCD algorithms are local and, for typical

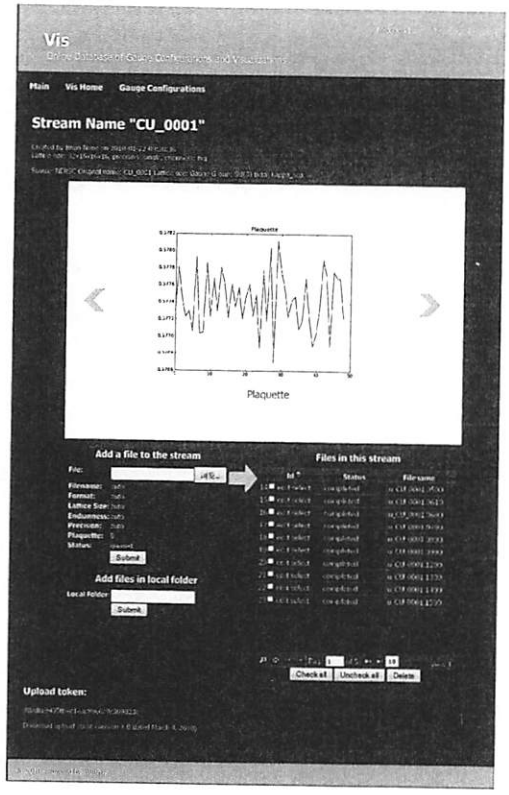

Figure 2.1: Screenshot of the stream view, including the MC history for the average plaquette and a list of files in the stream.

Figure 2.2: Screenshot of the processing view, showing information about the processing of files in a stream, including the status of the computations and the algorithm that was used.

production computations, they do not significantly change the total topological charge. The question, therefore, is whether or not they change the local topological charge density. If they do not, then the computation is biased because the MCMC must get stuck in a topological sector and is not sampling properly. Visualization techniques can be applied in this case, because we have, for example, been able to use them to show that the answer is yes – typical production computations are in fact not stuck in a topological sector.

Our goal is to automate the workflow of physicists working with this data and allow them to:

- store and share gauge configuration for multiple MCMC streams.
- schedule computing jobs for each stream, in particular the computation of the topological charge density.
- visualize the topological charge density (and other derived fields) using isosurfaces and/or volume plots.
- interact with the data using a web interface (rotate the topological charge and change visualization parameters).

2.2 Implementation

Vis, at its core, is a web application for storing collections of datasets and scheduling computations on the files in the sets. A set here is a MCMC stream, and the files in the set are the gauge configurations. Computations can be numerical algorithms and/or visualizations. The computations are submitted via an available Portable Batch System installation and can be parallel jobs.

Users can create an account in the system, login, and perform operations such as creating a new stream, uploading files into the stream, scheduling computations, searching and downloading streams submitted by other users, and viewing the results of computations performed on streams already in the system.

Some computations are scheduled automatically when data is uploaded, because new files need to be explored in order to detect their structure, they must be converted to a standard format, and then analyzed to extract some basic physical parameters that are important for cataloging the file and detecting possible errors.

Individual files can be very large (100M-1GB each) and therefore it may not be practical to upload them via the web interface, which does not support pausing and resuming. To avoid this problem, the system provides an alternate upload mechanism. When a new stream is created, a security token is issued to the user. The user can utilize a provided program to automatically upload every file from a local folder, authenticating via the downloaded token.

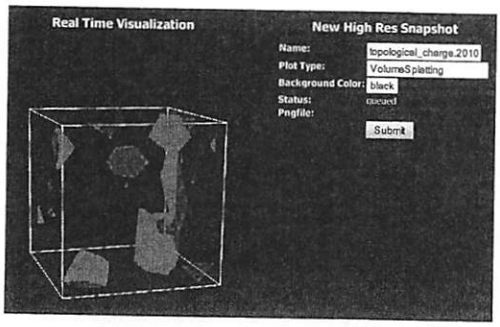

Figure 2.3: Screenshot showing widget that allows limited manipulation of visualized datasets.

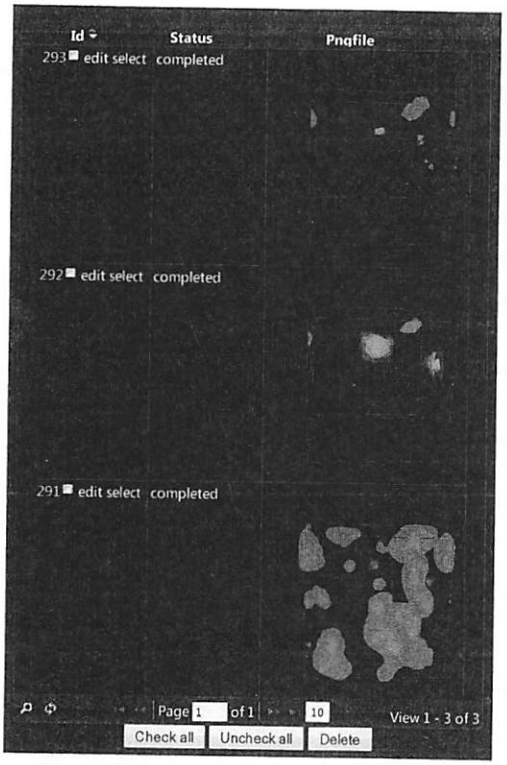

Figure 2.4: Using the browser-based interface, the user can set parameters for the visualization and then submit the job for processing.

The algorithm that computes the topological charge density generates output in VTK format. These files can, to a limited extent, be manipulated via the browser using a JavaScript widget that displays isosurfaces at 60% of the dataset's minimum and maximum values.

The user can choose a visualization angle, specify additional parameters, and then schedule a full visualization job to generate a high-res image.

The web interface was built using web2py [Di Pierro, 2010]. We utilize matplotlib [Hunter et al., 2010] for 2D plotting and VisIt for 3D visualization of volume plots and iso-surfaces, although we are exploring the possibility of moving to Vish [Benger, 2010] for the latter.

2.3 Summary

At this point Vis is primarily in the prototype stage, because it is hosted on a small PC and lacks the computing resources and bandwidth to transfer and store very large files. However, the program is fully functional and has been used to process some streams of gauge configurations that are made freely available by various groups via the NERSC archive [U.S. DOE, 2010].

Visualization algorithms can help physicists gain new insights into the physics of QCD, and Lattice QCD computations in particular. Our hope is that Vis can lower the barrier of entry, and enable physicists to look more deeply into their data. Vis can be downloaded from: `https://launchpad.net/qcdvis`

Acknowledgments

Project funded by the Department of Energy grant DEFC02-06ER41441.

Bibliography

[Altarelli, 2002] Altarelli, G. (2002). A QCD primer. URL: http://arxiv.org/abs/hep-ph/0204179v1.

[Benger, 2010] Benger, W. (2010). Vish website. URL: http://vish.origo.ethz.ch/.

[Di Pierro, 2010] Di Pierro, M. (2010). web2py website. URL: http://web2py.com.

[Hunter et al., 2010] Hunter, J., Dale, D., & Droettboom, M. (2010). matplotlib website. URL: http://matplotlib.sourceforge.net/.

[Massimo Di Pierro, 2006] Massimo Di Pierro (2006). AN ALGORITHMIC APPROACH TO QUANTUM FIELD THEORY. *International Journal of Modern Physics A*, 21, 405–448.

[Novaes, 1999] Novaes, S. F. (1999). Standard model: An Introduction. *Proceedings of the X J. A. Swieca Summer School.* URL: http://arxiv.org/abs/hep-ph/0001283v1.

[U.S. DOE, 2010] U.S. DOE (2010). Nersc website. URL: http://qcd.nersc.gov/.