Chapter 11
Parallel Quantum Chemistry at the Crossroads

Hubertus J. J. van Dam
Pacific Northwest National Laboratory, USA

ABSTRACT

Quantum chemistry was a compute intensive field from the beginning. It was also an early adopter of parallel computing, and hence, has more than twenty years of experience with parallelism. However, recently parallel computing has seen dramatic changes, such as the rise of multi-core architectures, hybrid computing, and the prospect of exa-scale machines requiring 1 billion concurrent threads. It is doubtful that current approaches can address the challenges ahead. As a result, the field finds itself at a crossroads, facing the challenge to successfully identify the way forward. This chapter tells a story in two parts. First, the achievements to date are considered, offering insights learned so far. Second, we look at paradigms based on directed acyclic graphs (DAG). The computer science community is strongly advocating this paradigm, but the quantum chemistry community has no experience with this approach. Therefore recent developments in that area will be discussed and their suitability for future parallel quantum chemistry computing demands considered.

INTRODUCTION

Quantum chemistry was formulated with the Schrödinger equation in1926 as the science that seeks to explain chemistry, chemical processes and properties of molecules from solutions of this

equation. As a result this field of science was a compute intensive enterprise from the start. Since the beginning, when the word "computer" referred to a job description rather than a machine (Kopplin, 2002), there have been dramatic developments both on the theory side and as well as the practical side. On the theory side the formulation of the problem in a basis set, the formulation of

theories to deal with the complex electron-electron interaction, such as Density Functional Theory (DFT) (Hohenberg & Kohn, 1964), Many Body Perturbation Theory (MBPT) (Møller & Plesset, 1934), and Coupled Cluster (CC) (Čížek, 1966), provided important handles on the problem. On the practical side the rapid development of computer technology and the design of efficient algorithms delivered access to the amount of computation needed to deliver useful answers.

In particular the development of computers followed initially a straightforward path of ever increasing compute power from single processing units. Most recently this trend has been broken as practical limits were hit. As it became impossible to simply crank up the clock speed even higher processor engineers opted instead for building chips with multiple processing units, the so called multi-core processors. This is a significant change. Although parallelism, as we currently know it, was introduced in quantum chemistry as early as 1988 (Clementi et al.,1988) and continuous and significant investment in parallel codes was made there was always the excuse that most machines were actually single processor machines to avoid becoming embroiled in code parallelization. However, now that every new machine is effectively a parallel machine it is clear that only parallel applications can utilize the available hardware effectively. But the changes go far beyond this point. Alongside the development of multi-core processors the development of Graphics Processing Units (GPU) towards General Purpose computing (GPGPU) created a new level of parallelism (Ufimtsev et al., 2008) (Vogt et al., 2008) (Genovese et al., 2009). The GPGPU's of today offer the potential of using 240 cores concurrently within a single physical machine. This is paired with the development of off-the-shelf solutions to connect multiple machines to build compute clusters. As a result relatively powerful machines can be built for low costs at present.

Nevertheless, the development towards exa-scale computers is likely going to be disruptive.

The reason for this lies in the energy requirements for such a machine. Having that the performance is specified and thereby the CPU power consumption to a large extent fixed it will be essential to save on memory and network power consumption to deliver these machines. This, however, raises the bar for writing scalable applications as there will be less space to keep data but it will also become more expensive to move data.

Against this back-drop of computer developments let us consider how close we have come to delivering a dream that was once captured in the title of a Europort2 project "Interactive Molecular Modeling through Parallelism" (Colbrook et al., 1995). What has been learned about parallelism in quantum chemistry and do we currently have a way to tackle the next level of parallelism?

With respect to answering these questions we find the field currently at a crossroads. Over the next decade machines will be developed that are a thousand times more powerful than the top end machines of today. These machines can be developed only by increasing the concurrency. However, many believe that the software development paradigms adopted so far cannot be stretched to such high levels of concurrency. Instead, computer scientists propose to write codes using directed acyclic graphs (DAG) that make the data dependencies explicit. This, they claim, allows parallelism to be exploited much more flexibly. In quantum chemistry there currently is no experience with paradigms of this kind. Hence, the field is faced with the question in which direction to proceed. In this chapter the way parallelism has been approached to date will be considered first. This provides insight into what has been achieved, how parallelism has been addressed, and what has been learned along the way. Next, looking forward, DAG based approaches are considered. In particular a number of existing frameworks and programming models are discussed to provide a flavor of the direction in which parallel programming may develop in the years to come.

Related Content

Topology-Aware Load-Balance Schemes for Heterogeneous Graph Processing
 (2018). *Creativity in Load-Balance Schemes for Multi/Many-Core Heterogeneous Graph Computing:
Emerging Research and Opportunities (pp. 113-143).*
www.irma-international.org/chapter/topology-aware-load-balance-schemes-for-heterogeneous-graph-processing/195894

Implementation of FFT on General-Purpose Architectures for FPGA
Fabio Garzia, Roberto Airoldiand Jari Nurmi (2012). *Computer Engineering: Concepts, Methodologies,
Tools and Applications (pp. 658-676).*
www.irma-international.org/chapter/implementation-fft-general-purpose-architectures/62470

Whale Optimization Algorithm With Wavelet Mutation for the Solution of Optimal Power Flow
Problem
V. Mukherjee, Aparajita Mukherjeeand Dharmbir Prasad (2018). *Handbook of Research on Predictive
Modeling and Optimization Methods in Science and Engineering (pp. 500-553).*
www.irma-international.org/chapter/whale-optimization-algorithm-with-wavelet-mutation-for-the-solution-of-optimal-
power-flow-problem/206764

True Color Image Segmentation by MUSIG Activation Function Using Self-Supervised
QMLSONN Architecture With Context-Sensitive Thresholding
Pankaj Paland Siddhartha Bhattacharyya (2018). *Quantum-Inspired Intelligent Systems for Multimedia
Data Analysis (pp. 213-261).*
www.irma-international.org/chapter/true-color-image-segmentation-by-musig-activation-function-using-self-supervised-
qmlsonn-architecture-with-context-sensitive-thresholding/202549

Attaining Semantic Enterprise Interoperability Through Ontology Architectural Patterns
Rishi Kanth Saripalleand Steven A. Demurjian (2018). *Computer Systems and Software Engineering:
Concepts, Methodologies, Tools, and Applications (pp. 705-740).*
www.irma-international.org/chapter/attaining-semantic-enterprise-interoperability-through-ontology-architectural-
patterns/192899