Heterogeneous Hardware Acceleration of Parallel Algorithms

Ra Inta, David J. Bowman and Susan M. Scott

Department of Quantum Science,
The Australian National University, Canberra ACT 0200, Australia

000: Introduction

In recent years, there has been a growing interest in using heterogeneous computing systems to solve complex problems. These systems combine multiple processors with different architectures to leverage their strengths and overcome their limitations. The use of GPUs (Graphics Processing Units), FPGAs (Field-Programmable Gate Arrays), and CPUs (Central Processing Units) in parallel computing has become increasingly popular due to their ability to handle diverse computational tasks efficiently.

The initial concept was to mediate separate clusters of CPUs, GPUs, and FPGAs (Figure 1). However, there are many other data-hungry systems about to come on-line, such as the Large Synoptic Survey Telescope (Figure 1). We are experiencing an ever-increasing deluge of data from large scientific projects, requiring commensurately significant computational power.

Because of the three hardware classes, we named the system the Chimera Computing System (Figure 1). Excellent at dense linear algebra, GPUs are particularly well-suited for applications that require massive parallelism. FPGAs, on the other hand, are programmable, offering the advantage of customizing the hardware to specific applications. CPUs, with their rich set of instructions, are best suited for sequential processing and control logic.

001: Advantages of Hardware Acceleration

<table>
<thead>
<tr>
<th>Platform</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>Analysis workflow, multithreading, high peak performance, limited flexibility</td>
<td>Processor hungry, limited parallelism, highly rigid instruction set</td>
</tr>
<tr>
<td>GPU</td>
<td>Unrestricted scalability, flexibility, low power consumption, limited instruction set</td>
<td>Expensive, specialized instruction set, prohibitive development time</td>
</tr>
<tr>
<td>FPGA</td>
<td>Ultra-fine granularity, flexibility, programming freedom, low power consumption</td>
<td>Expensive, specialized instruction set, prohibitive development time</td>
</tr>
</tbody>
</table>

Table 1: Comparison of the advantages and disadvantages of CPU-based calculations to those of the GPU and FPGA. This list is not comprehensive, nor is it exhaustive.

010: The ‘Chimera’ Computing System

In the Chimera Computing System, there are two main compute nodes: the CPU and the GPU. The CPU handles the general-purpose computations, while the GPU is optimized for high-speed, parallel processing. This hybrid system is a powerful tool for complex scientific simulations and data analysis.

The Chimera Computing System is designed to be flexible and scalable. The number of compute nodes can be increased or decreased according to the needs of the application. This system has been successfully used in various research domains, including astrophysics, climate modeling, and molecular dynamics.

011: Performance

Performance on the type of heterogeneous platforms is highly algorithmic dependent. One of the simple algorithms that illustrates the concept of this platform is the parallel implementation of the Fast Fourier Transform (Figure 1). The number of floating-point operations is significantly reduced compared to the serial implementation.

010: Analysis of Parallel Architectures via the ‘Thirteen Dwarves of Berkeley’

Table 2: The “Thirteen Dwarves of Berkeley.” This is a list of the main classes or categories of parallel machines. The classification is based on the level of hardware and software independence.

<table>
<thead>
<tr>
<th>Dwarves</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workstations</td>
<td>Minimal hardware and software independence</td>
</tr>
<tr>
<td>Workstations</td>
<td>Partial hardware and software independence</td>
</tr>
<tr>
<td>Workstations</td>
<td>Complete hardware and software independence</td>
</tr>
</tbody>
</table>

We have developed a comprehensive model that predicts the performance of the system for different algorithms. The model takes into account the hardware characteristics and the algorithms’ requirements. This system is a valuable tool for designing and optimizing parallel systems.

101: Conclusion

The Chimera Computing System is a versatile and powerful tool for parallel computing. It can be used in a wide range of applications, from scientific simulations to data analysis. The flexibility of this system allows it to be adapted to the specific needs of each application.

References


Contact

R. Inta, D. Bowman, S. M. Scott
E-mail: raint@anu.edu.au

Australian National University
Canberra ACT 0200, Australia