The CUDA Handbook: A Comprehensive Guide to GPU Programming
The CUDA Handbook begins where CUDA by Example (Addison-Wesley, 2011) leaves off, discussing CUDA hardware and software in greater detail and covering both CUDA 5.0 and Kepler. Every CUDA developer, from the casual to the most sophisticated, will find something here of interest and immediate usefulness. Newer CUDA developers will see how the hardware processes commands and how the driver checks progress; more experienced CUDA developers will appreciate the expert coverage of topics such as the driver API and context migration, as well as the guidance on how best to structure CPU/GPU data interchange and synchronization. The accompanying open source code “more than 25,000 lines of it, freely available at www.cudahandbook.com” “is specifically intended to be reused and repurposed by developers. Designed to be both a comprehensive reference and a practical cookbook, the text is divided into the following three parts: Part I, Overview, gives high-level descriptions of the hardware and software that make CUDA possible. Part II, Details, provides thorough descriptions of every aspect of CUDA, including Memory Streams and events Models of execution, including the dynamic parallelism feature, new with CUDA 5.0 and SM 3.5 The streaming multiprocessors, including descriptions of all features through SM 3.5 Programming multiple GPUs Texturing The source code accompanying Part II is presented as reusable microbenchmarks and microdemos, designed to expose specific hardware characteristics or highlight specific use cases. Part III, Select Applications, details specific families of CUDA applications and key parallel algorithms, including Streaming workloads Reduction Parallel prefix sum (Scan) N-body Image Processing These algorithms cover the full range of potential CUDA applications.
"The CUDA Handbook" is the largest and latest (June 2013) of NVIDIA's series of GPU programming books. It is also the most comprehensive and useful GPU programming reference for programmers to date. It's a tough world out there for programmers who are trying to keep up with changes in technology and this reference makes the future a much more comfortable place to live. Learn about GPGPU programming and get ahead of the crowd. For those programmers who haven't had the time to perceive the changes, GPU programming is a current change in programming design that is sweeping the world of network VOIP management, parallel analysis and simulation, and even supercomputing in a single box. I have personally run a Starfield Simulation on a portable with an i7 processor that increased in speed 112 times by using the internal NVIDIA GeForce 570M. The Starfield frame time reduced from about 2 seconds to about .015 sec. Imagine what I could do with a GeForce 690! Charts indicate that it might exceed 700 times the computing speed!!This book not only tells me how to arrange the software to work with the NVIDIA SDK, but it also shows me the important differences in the architecture of many of the NVIDIA cards to obtain optimum performance. The world of computing is still filled with 32 bit machines (or OS systems) using most of their memory to get their assigned tasks completed. Many of these machines do not have even four core CPUs, forget having over 4GB of memory. They fill computers in production devices, desktops in database support companies, and the racks of IT departments everywhere. The need for faster and more computing does not slow down or stop for these hardware limits. Ant the cost to replace them outright is prohibitive.

Download to continue reading...