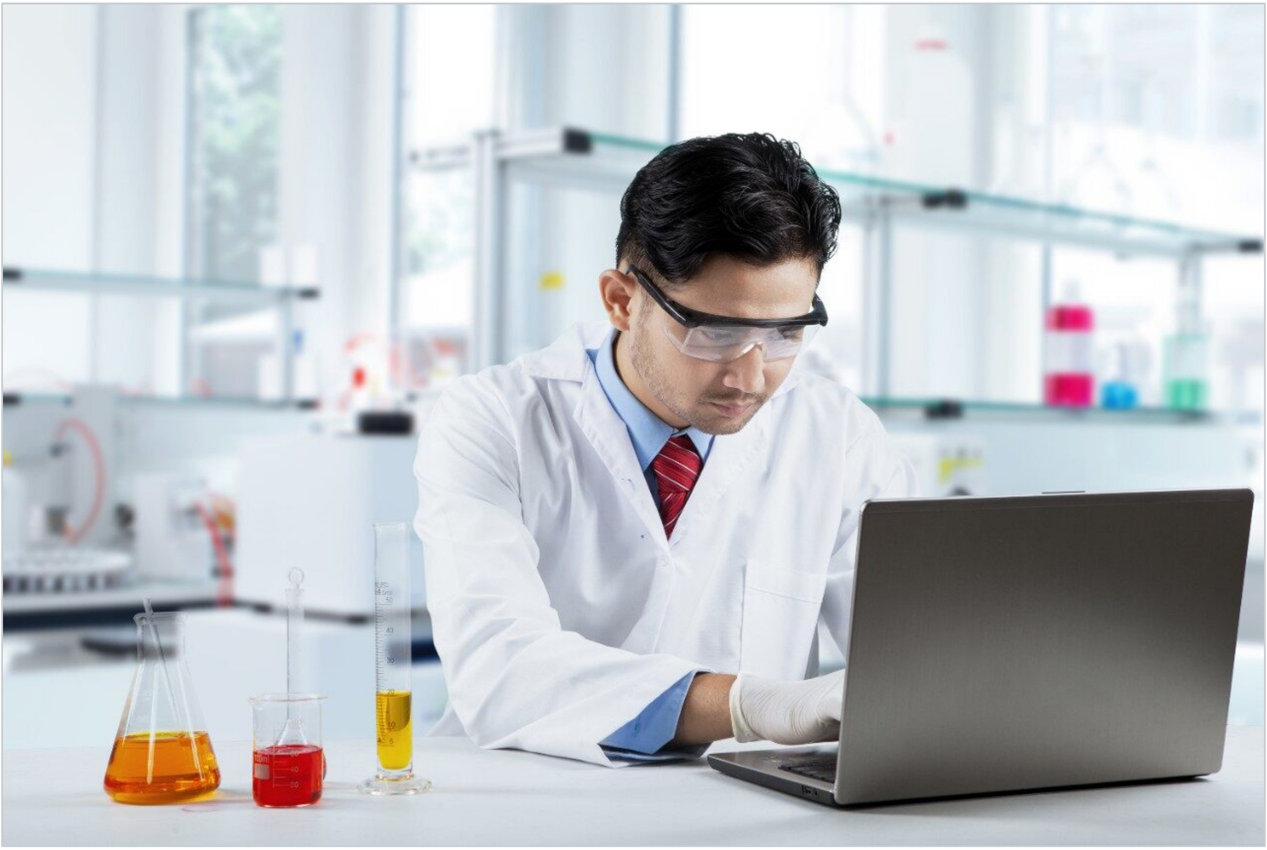
Applikationsbericht

The Implementation of High- Performance GPU Computing to Revolutionize LC-MS^E Data Processing

Waters Corporation



This is an Application Brief and does not contain a detailed Experimental section.

Abstract

The use of the multiplexed data acquisition strategy, LC-MS^E provides the most comprehensive and unbiased method of analyzing complex biological systems. However, it is computationally intensive and places a heavy demand on informatics post-processing. Here we describe a novel solution to significantly speed up associated peak detection and data processing times.

Benefits

Implementation of LC-MS^E on a GPU reduces computation time and speeds up overall processing time by a minimum of 5x.

Introduction

LC-MS^E is a key method of data acquisition that utilizes parallel sequencing of all detectable peptides as they elute from the UPLC Column with the use of high mass resolution and accuracy, which helps maximize coverage and confidence in results. When analyzing and utilizing this data, it is imperative to uniquely identify each ion in both the low and elevated energy functions, and then to correlate the eluting peptides with their associated fragments, using retention time alignment. This provides highly specific and sensitive information for database searching and quantitative profiling.

The processing of these large-scale datasets requires significant computational power and is often a bottleneck in determining the constituent peptides and proteins. In this technical brief, we describe the implementation of peak detection and data processing algorithms on a Graphics Processing Unit (GPU) that resulted in significant time savings for processing LC-MS^E data.

The GPU processes all graphics and images shown in a computer display. In this sense, it is a dedicated processor, whereas the CPU (Central Processing Unit) is a general processor that handles all other tasks in a computer. In recent years, the GPU's computation power has increased at a higher rate than the CPU, mostly driven by the need for higher quality graphics in today's consumer and professional applications. Because of its greater computation power, the GPU is now being used for tasks other than graphics. For example, in the last three years, it has been used for computationally intensive applications in various scientific and engineering disciplines.

Unlike the CPU, which is a series processor that can only process one program instruction at a time, the GPU is a parallel processor that can process many program instructions in parallel. Another major difference is the number of cores that can be placed in a chip. Today, the CPU packs up to six cores in a chip, whereas the GPU packs up to 448.

The GPU achieves significant reductions in computation time in compute-intense applications that can be parallelized, such as applications that do the same repetitive computation on different data.

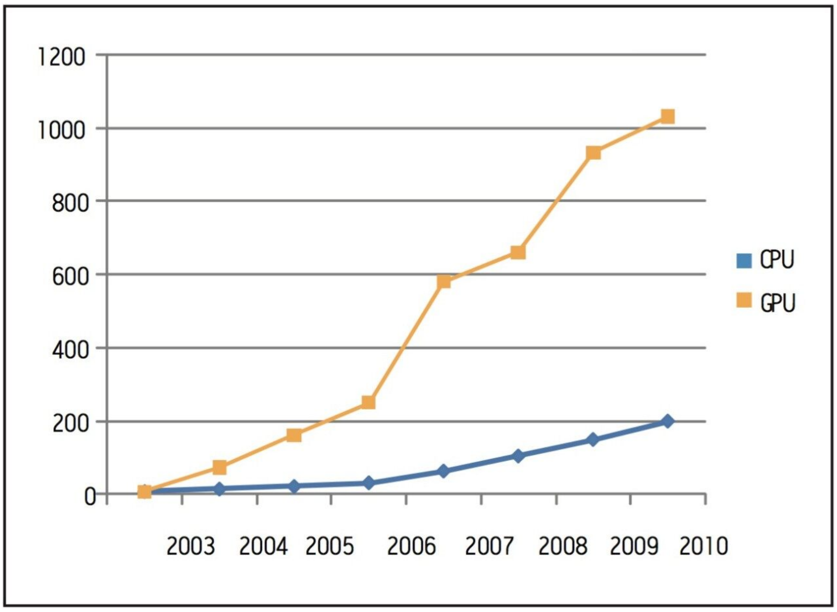


Figure 1. Historical comparison of the computation power difference between the CPU and GPU in GFLOPS.

Results and Discussion

In LC-MS^E data processing, peak detection from the raw data is performed by Apex 3D, an algorithm which has been previously described in detail[1]. By running this algorithm on the GPU, we obtained an 50 x increase in speed in the computational sections of the program for a typical data size. With this reduced computation time, the data access sections of the program became dominant in the overall processing time, which resulted in a net speedup of approx 5x.

This effect is less important as data become larger and more complex, as shown in Figures 2 and 3. In these cases, the time in data access is less relevant and the net speedup is closer to the computational speedup.

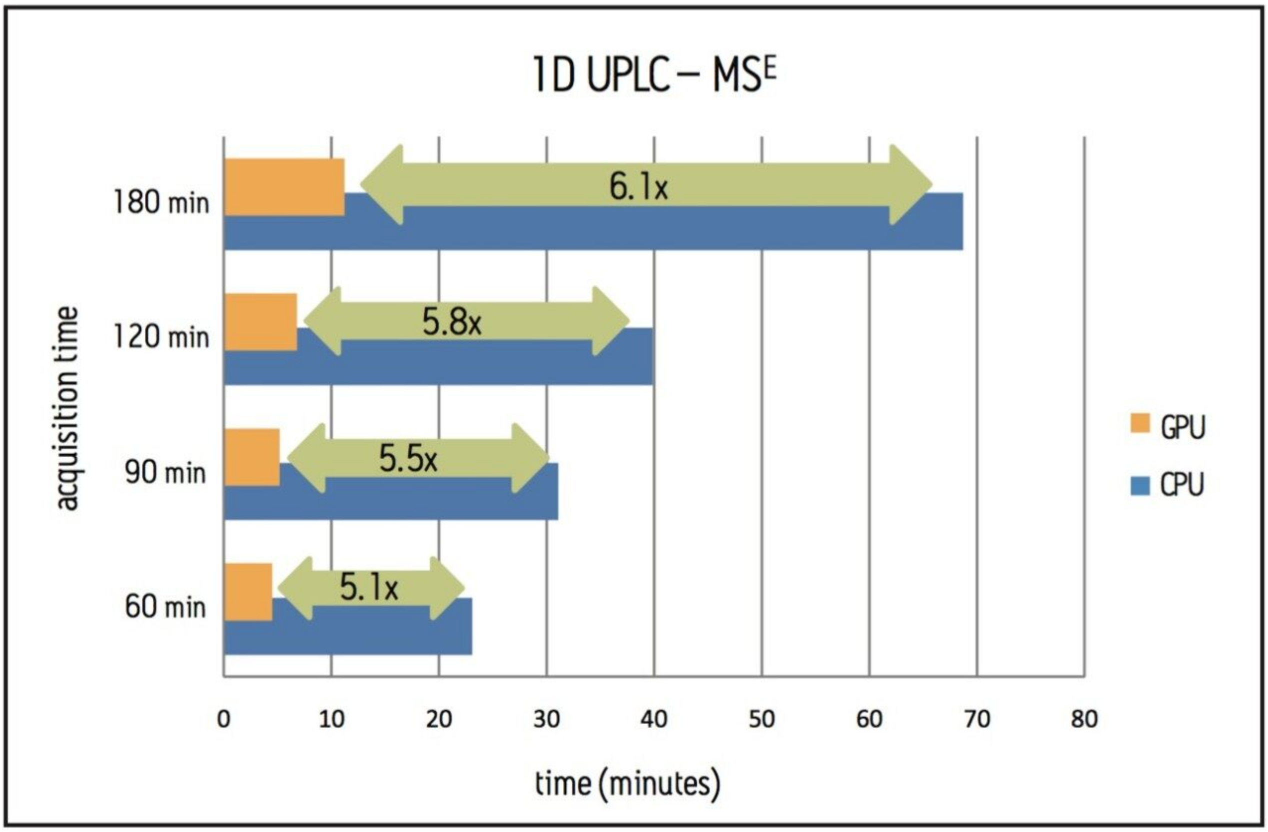


Figure 2. Improvement in performance using a 240-core GPU compared with a quad-core CPU for processing LC-MS^E data files of varying file size from different chromatographic separations.

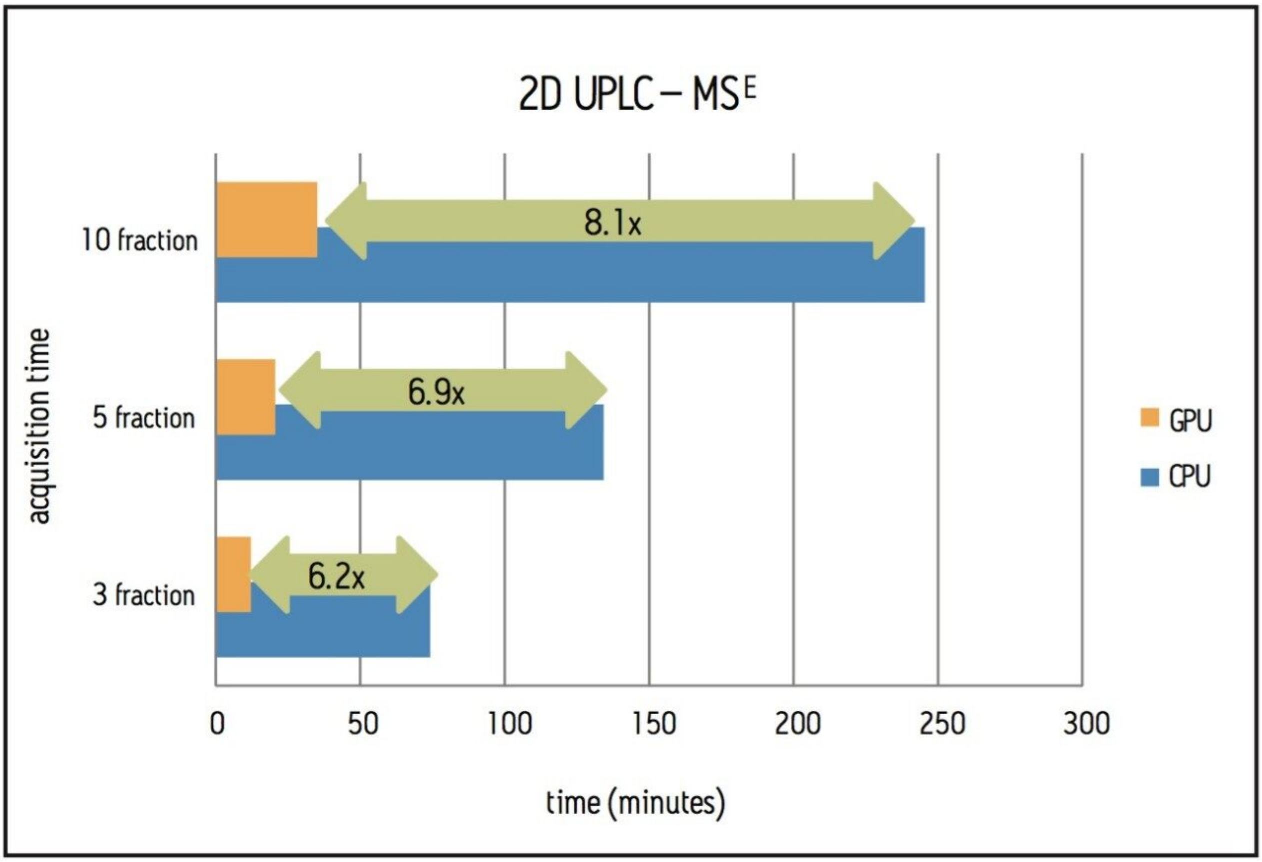


Figure 3. Comparison of data processing times for LC-MS^E data files from 3-, 5-, and 10-step 2D-LC experiments using a 240-core GPU compared with a quad-core CPU.

Conclusion

Here we have described the implementation of LC-MS^E data algorithms on a GPU that significantly enhances the speed of data processing. This will be crucial in processing multi-dimensional LC-IMS-MS^E data in the near future.

Featured Products