A variety of research exists for the processing of continuous queries in large, mobile environments. Each method tries, in its own way, to address the computational bottleneck of constantly processing so many queries. For this research, we present a two-pronged approach at addressing this problem. Firstly, we introduce an efficient and scalable system for monitoring traditional, continuous queries by leveraging the parallel processing capability of the Graphics Processing Unit. We examine a naive CPU-based solution for continuous range-monitoring queries, and we then extend this system using the GPU. Additionally, with mobile communication devices becoming commodity, location-based services will become ubiquitous. To cope with the very high intensity of location-based queries, we propose a view oriented approach of the location database, thereby reducing computation costs by exploiting computation sharing amongst queries requiring the same view. Our studies show that by exploiting the parallel processing power of the GPU, we are able to significantly scale the number of mobile objects, while maintaining an acceptable level of performance.

Our second approach was to view this research problem as one belonging to the domain of data streams. For example, the output of a GPS receiver, monitoring the position of a mobile object, is viewed as a data stream of location updates. This data stream of location updates, along with those from the plausibly many other mobile objects, is received at a centralized server, which processes the streams upon arrival, effectively updating the answers to the currently active queries in real time.

For this second approach, we present GEDS, a scalable, Graphics Processing Unit (GPU)-based framework for the evaluation of continuous spatio-temporal queries over spatio-temporal data streams. Specifically, GEDS employs the computation sharing and parallel processing paradigms to deliver scalability in the evaluation of continuous, spatio-temporal range queries and continuous, spatio-temporal kNN queries. The GEDS framework utilizes the parallel processing capability of the GPU, a stream processor by trade, to handle the computation required in this application. Experimental evaluation shows promising performance and shows the scalability and efficacy of GEDS in spatio-temporal data streaming environments. Additional performance studies demonstrate that, even in light of the costs associated with memory transfers, the parallel processing power provided by GEDS clearly counters and outweighs any associated costs.

Finally, in an effort to move beyond the analysis of specific algorithms over the GEDS framework, we take a broader approach in our analysis of GPU computing. What algorithms are appropriate for the GPU? What types of applications can benefit from the parallel and stream processing power of the GPU? And can we identify a class of algorithms that are best suited for GPU computing? To answer these questions, we develop an abstract performance model, detailing the relationship between the CPU and the GPU. From this model, we are able to extrapolate a list of attributes common to successful GPU-based applications, thereby providing insight into which algorithms and applications are best suited for the GPU and also providing an estimated theoretical speedup for said GPU-based applications.

Major: Computer Science

Educational Career:
Bachelor’s of Computer Science, BS, 2006, University of Central Florida
Master’s of Computer Science, MS, 2009, University of Central Florida

Committee in Charge:
Prof. Ratan Guha, Chair, Department of Electrical Engineering and Computer Science
Prof. Mostafa Bassiouni, UCF / Department of Electrical Engineering and Computer Science
Approved for distribution by Prof. Ratan Guha, Committee Chair, on March 12, 2012.

The public is welcome to attend.