

UNIVERSITY OF SOUTH FLORIDA

Defense of a Doctoral Dissertation

System Support of Concurrent Database Query Processing on a GPU

by

Hao Li

For the Ph.D. degree in Computer Science and Engineering

The unrivaled computing capabilities of modern GPUs meet the demand of processing massive amounts of data seen in many application domains. We propose a GPU-based DBMS (G-DBMS) that can run multiple tasks concurrently. To that end, system-level management mechanisms such as resource allocation and buffer manager are needed to build such a concurrent database query processing system and fully unleash the GPUs' computing power. However, CUDA does not provide enough OS-level functionalities to support it. Thus, our research is focusing on implementing the optimization of resource allocation and GPU memory management. Firstly, we have explored the single compute-bound kernel modeling on GPUs under NVidia's CUDA framework and provide in-depth anatomy of NVidia's concurrent kernel execution mechanism (CUDA stream), which is the foundation of the resource allocation in CUDA. Second, we study resource allocation of multiple GPU applications towards optimization of system throughput in the context of systems. Lastly, we develop a novel Buffer Manager on GPU that represents the buffer in a bitmap contained metadata including mutex of each page. Upon them, we are able to build the prototype of G-DBMS.

Examining Committee

Ming Ji, Ph.D., Chairperson
Yicheng Tu, Ph.D., Major Professor
Adriana Iamnitchi, Ph.D.
Bo Zeng, Ph.D.
Kandethody Ramachandran, Ph.D.
Lingling Fan, Ph.D.
Srinivas Katkoori, Ph.D.

Tuesday November 10, 2020
11:00 AM

Online (Collaborate Ultra)
Please email for more information
haoli1@usf.edu
THE PUBLIC IS INVITED

Publications

- 1) **Hao Li**, Di Yu, Anand Kumar, and Yicheng Tu. "Performance modeling in CUDA streams—A means for high-throughput data processing." In 2014 IEEE International Conference on Big Data (Big Data), pp. 301-310. IEEE, 2014.
- 2) Ran Rui, **Hao Li**, and Yicheng Tu. "Join algorithms on GPUs: A revisit after seven years." In 2015 IEEE International Conference on Big Data (Big Data), pp. 2541-2550. IEEE, 2015.
- 3) **Hao Li**, Yicheng Tu, and Bo Zeng. "Concurrent query processing in a GPU-based database system." PloS one 14, no. 4 (2019): e0214720.

Robert Bishop, Ph.D.
Dean, College of Engineering

Dwayne Smith, Ph.D.
Dean, Office of Graduate Studies

Disability Accommodations:

If you require a reasonable accommodation to participate, please contact the Office of Diversity & Equal Opportunity at 813-974-4373 at least five (5) working days prior to the event.