Accelerating Graph Applications on Parallel Heterogeneous Architectures

How will computing power increase in the future? We need to use parallelism more effectively.

Innovative parallel system designs like GPUs and the TPU have achieved significant performance improvements. However, a closer look will reveal that these gains have largely been achieved on “dense” computations, in which data is accessed in regular, predictable patterns. But we also need to achieve improvements in “sparse” graph computations, because these computations are a natural way to represent and analyze important data structures, such as social networks and consumer behaviors.

This talk will present some of the current work at Princeton on parallel heterogeneous systems – and how to use them to accelerate graph applications. One part of this effort has aimed at optimizing graph applications for a prevalent existing system: the CPU/GPU SoC. A domain specific language and an optimizing compiler have been created to conduct a large empirical study on a range of GPU systems. This talk will present some key results of the study, including significant speedups, horrible slowdowns, and a new robust definition of performance portability.

Another part of the research program is the design of a new parallel heterogeneous architecture, with reconfigurable and programmable communication structures tailored for graph computations. This system addresses key memory bottlenecks by pairing up the simple cores of a many-core system in a producer/consumer relationship. Simulation results show that our design is significantly more efficient (up to 20x) than existing processors.

Dr. Tyler Sorensen is an assistant professor of Computer Science at University of California Santa Barbara. Tyler has recently been a postdoctoral researcher at Princeton University working with Professor Margaret Martonosi. He is involved in the DECADES project, a large collaboration (funded by DARPA’s Software-Defined Hardware program) that is designing new heterogeneous chips to accelerate emerging machine learning and graph analytic workloads. Tyler has been largely involved in the compiler, runtime, and data-supply designs of the system.

Tyler has a Ph.D. from Imperial College London, where we worked on programming models for GPUs. His work has received several distinguished paper awards at top programming language and software engineering conferences (PLDI, IISWC and FSE). Tyler is an invited academic member of the Khronos Group, a non-profit industry consortium that defines heterogeneous programming language standards.

Date: Thursday, November 19, 2020, 8:00pm
Place: ONLINE MEETING – registration required
How to register:
• Send email to PrincetonACM@gmail.com
• OR Register on Meetup.com
  (http://meetup.com/IEEE-Princeton-Central-Jersey-Section)
Information: Dennis Mancl (908) 285-1066
On-line info: http://PrincetonACM.acm.org

All Princeton ACM / IEEE-CS meetings for fall/winter 2020-21 will be held “on-line”. When you register for the meeting, you will receive an email with instructions for how to connect to the talk.

All Princeton ACM / IEEE-CS meetings are open to the public. Students and their parents are welcome. There is no admission charge.