Data-Intensive Super Computing: Taking Google-Style Computing Beyond Web Search

By
Dr. Randy Bryant (CMU)

Date: **March 10th, Friday.**
Time: **2:30 PM**
Venue: **155 Coates Hall**

Abstract

Web search engines have become fixtures in our society, but few people realize that they are actually publicly accessible supercomputing systems, where a single query can unleash the power of several hundred processors operating on a data set of over 200 terabytes. With Internet search, computing has risen to entirely new levels of scale, especially in terms of the sizes of the data sets involved. Google and its competitors have created a new class of large-scale computer systems, which we label "Data-Intensive Super Computer" (DISC) systems. DISC systems differ from conventional supercomputers in their focus is on data: they acquire and maintain continually changing data sets, in addition to performing large-scale computations over the data.

With the massive amounts of data arising from such diverse sources as telescope imagery, medical records, online transaction records, and web pages, DISC systems have the potential to achieve major advances in science, health care, business, and information access. DISC opens up many important research topics in system design, resource management, programming models, parallel algorithms, and applications. By engaging the academic research community in these issues, we can more systematically and in a more open forum explore fundamental aspects of a societally important style of computing.

BIO:

Randal E. Bryant is Dean of the Carnegie Mellon University School of Computer Science. He has been on the faculty at Carnegie Mellon for 23 years, starting as an Assistant Professor and progressing to his current rank of University Professor.

Much of Dr. Bryant's research focuses on methods for formally verifying digital hardware, and more recently some forms of software. His 1986 paper on symbolic Boolean manipulation using Ordered Binary Decision Diagrams (BDDs) has the highest citation count of any publication in the Citeseer database of computer science literature. In addition, he has developed several techniques to verify

circuits by symbolic simulation, with levels of abstraction ranging from transistors to very high-level representations. Most recently, he has been investigating ways that large-scale computer systems can be organized and programmed to solve data-intensive problems.

Dr. Bryant is a fellow of the IEEE and the ACM, as well as a member of the National Academy of Engineering. His awards include: the 2007 IEEE Emmanuel Piore Award, the 1997 ACM Kanellakis Theory and Practice Award (shared with Edmund M. Clarke, Ken McMillan, and Allen Emerson) for contributing to the development of symbolic model checking, and the 1989 IEEE W.R.G. Baker Prize for the best paper appearing in any IEEE publication during the preceding year.

Dr. Bryant received his B.S. in Applied Mathematics from the University of Michigan in 1973, and his PhD from MIT in 1981. He was on the faculty at Caltech from 1981 to 1984.

The IT Eminent Lecture Series (ITELS) is jointly sponsored by LSU-CCT and LSU Computer Science Department.