

Dr. Randy Bryant

Data-Intensive Scalable Computing

Abstract

Web search engines have become fixtures in our society, but few people realize that they are actually publicly accessible, massive computing 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 Scalable 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. DISC points the way to new ways of organizing large-scale computing systems to be more robust, scalable, and cost effective than are current high-performance computing systems.

Programs for DISC systems must be written in ways that allows them to be executed in a loosely-coupled asynchronous environment, such as the Map/Reduce framework pioneered by Google. Although Map/Reduce has surprisingly broad applicability, a richer set of programming languages and models is required to realize the full potential of DISC.

Bio

Randal E. Bryant is Dean of the Carnegie Mellon University School of Computer Science. He has been on the faculty at Carnegie Mellon since 1984, starting as an Assistant Professor and progressing to his current rank of University Professor.

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 one of the highest citation counts of any publication in the computer science literature. Along with David R. O'Hallaron, he authored the textbook "Computer Systems: A Programmer's Perspective", now used in over 130 universities worldwide and translated into Chinese and Russian.

Since 2007, Dr. Bryant has been advocating Data-Intensive Scalable Computing (DISC) as a research and educational priority for universities. This area builds on advances in hardware and software technology, the availability of massive data sources, such as the Internet and large-scale sensor systems, and advanced machine learning algorithms to extract important insights from very large data sets. This technology is already well established at companies such as Google, Yahoo!, Microsoft, and Amazon, and it promises to provide new capabilities for scientific research, business, medicine, and society.

Dr. Bryant is a fellow of the IEEE and the ACM, as well as a member of the National Academy of Engineering and the American Academy of Arts and Sciences. His awards include: the

2010 ACM/IEEE A. Richard Newton Award, the 2009 IEEE/EDAC Kaufman Award, 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.