Abstract:

After a brief review of how map-reduce (Hadoop, really) works, we shall look at the trade-off that needs to be made when designing map-reduce algorithms for problems that are not embarrassingly parallel. In particular, the less data that one reducer is able to handle, the greater the total amount of data that must be communicated from mappers to reducers. We can view this trade-off as a function that gives the "replication rate" (average number of copies of an input communicated from mappers to reducers) in terms of the "reducer size" (number of inputs that can be accommodated at a reducer). For some interesting problems, including matrix multiplication, finding triangles in a graph, and finding bit strings at Hamming distance 1, we can get precise lower bounds on this function, and also match the lower bounds with algorithms that achieve the minimum replication rate for a given reducer size.

Biography:

Jeff Ullman is the Stanford W. Ascherman Professor of Engineering (Emeritus) in the Department of Computer Science at Stanford and CEO of Gradiance Corp. He received the B.S. degree from Columbia University in 1963 and the PhD from Princeton in 1966. Prior to his appointment at Stanford in 1979, he was a member of the technical staff of Bell Laboratories from 1966-1969, and on the faculty of Princeton University between 1969 and 1979. From 1990-1994, he was chair of the Stanford Computer Science Department. Ullman was elected to the National Academy of Engineering in 1989, the American Academy of Arts and Sciences in 2012, and has held Guggenheim and Einstein Fellowships. He has received the Sigmod Contributions Award (1996), the ACM Karl V. Karlstrom Outstanding Educator Award (1998), the Knuth Prize (2000), the Sigmod E. F. Codd Innovations Award (2006), and the IEEE von Neumann Medal (2010). He is the author of 16 books, including books on database systems, compilers, automata theory, and algorithms.

Algorithm Design for Map-Reduce

Dr. Jeff Ullman
Professor Emeritus, Stanford University
Member National Academy of Engineering; ACM Fellow;
Knuth Prize; IEEE John von Neumann Medal; ACM Karl
Karlstrom Outstanding Educator Award
Graham Center Ballrooms
10:00am to 11:00am