Programming Big Data in X10

Prof. Vijay Saraswat[1]

[1] *IBM T. J. Watson research center*
Email: vijay@saraswat.org

Timetable: 12 hrs. First lecture on April 15, 2013, 09:00 (dates already fixed, see the calendar), Torre Archimede, Room 2BC/30

Course requirements: Java and/or C++, programming experience, introduction to concurrency

Examination and grading: Two in-class quizzes, one two-week programming assignment (code to be run on cluster)

SSD: INF/01 - Computer Science

Aim: To introduce students to programming with large data sets

Course contents:
This course will develop X10 as a programming model for scale-out computation with big data sets.

Topics covered: Introduction to Big Data and data analysis problems. Introduction to X10 and the APGAS (Asynchronous, Partitioned Global Address Space model). and basic multi-place programming idioms. Map Reduce programming – Hadoop and the X10 Map Reduce engine. Global Rules Engine – executing multiple, cooperating rule engines. Text analytics – sentiment analysis. Linear algebra – distributed matrix multiplication, matrix factorization. Graph analytics – between-ness centrality, k-clique. State-space search – unbalanced tree search problem. Extending R and Python with APGAS constructs for big data programming.