

Parallelizing the BLAST Algorithm: Feasible or Not?

The field of Bioinformatics Research, especially in the field of coding and classifying genes, has a need for fast string matching algorithms. In the Computer Science and Mathematics Division at the Oak Ridge National Laboratory (ORNL), High Performance Cluster (HPC) computing has been applied to many different areas, from computational biology to computational material science. The purpose of this project is to do a study on the Basic Local Alignment Search Tool (BLAST) algorithm: define the structure of the BLAST algorithm, state why the algorithm is valuable as a bioinformatics database tool, and explore the ways of increasing this algorithm's effectiveness and speed. BLAST stands for Basic Local Alignment Search Tool, and it is used in bioinformatics to find alignments between strings. BLAST is a heuristic algorithm that uses the technique of finding matches between fragments of a query string and a target database. This eliminates much of the data in a database without running a full comparison for each letter in the search string. Once query and database string alignments are found (if the fragments match within a certain threshold), the full strings are matched. Several methods for parallelizing BLAST have been explored and this information will be summarized in this paper. This paper will conclude with a number of potential methods for increasing the speed and effectiveness of BLAST.

Category: Computer Science and Mathematics

Student's Name:	William Burke
School Student Attends:	York College, C.U.N.Y.
Names(s) of Mentor(s)	Dr. Stephen Scott and John Mugler
Division:	Computer Science and Mathematics
Program:	Research Alliance for Minorities