

DNA

Deoxyribonucleic acid (DNA) is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms and some viruses. The main role of DNA molecules is the long-term storage of information . DNA is often compared to a set of blueprints or a recipe, or a code, since it contains the instructions needed to construct other components of cells . This project works based on the Grid architecture i.e. based on Client-Server architecture. The various tasks are provided by standing on this architecture. DNA operation done based on two methods named Finger print identification and Finding parent. This is carried out by pattern matching and matching percentage respectively. We have to create a database which is applicable to do a time consuming task on a short time period. The database is taken from the site mpiBlast. mpiBLAST is an open-source parallelization of BLAST that achieves super linear speed-up by segmenting a BLAST database and then having each node in a computational cluster search a unique portion of the database. Database segmentation permits each node to search a smaller portion of the database. Thus when the input is given the database checks for the Kilo Bytes and for a value greater than a particular value the result is obtained.

Grid computing:

Grid computing (or the use of a computational grid) is the application of several computers to a single problem at the same time - usually to a scientific or technical problem that requires a great number of computer processing cycles or access to large amounts of data.

Grid computing depends on software to divide and apportion pieces of a program among several computers, sometimes up to many thousands. Grid computing can also be thought of as distributed and large-scale cluster computing, as well as a form of network-distributed parallel processing. It can be small -- confined to a network of computer workstations within a corporation or it can be large -- a public collaboration across many companies or networks.

We will implement the Grid computing into this for enabling fast searching in the web.

This Project Has The Following Modules:

Server :

In this module it includes the Client status, Client menu and Task list. The server watches for the current status of the client in a network. Thus the needed situation can be achieved by the server. From the client menu the various tasks performed by the client can be obtained. From the task list further details about each task can be obtained.

Client :

In this module finger print identification and parent finding is performed. This done by pattern matching and the percentage to which the input is matched with the parent. Thus client selects the appropriate input from the whole and undergoes further operations.

Pattern matching :

In this module it deals with the checking for the presence of the constituents of a given pattern. In contrast to pattern recognition, the pattern is rigidly specified. Such a pattern concerns conventionally either sequences or tree structures . Pattern matching is used to test whether things have a desired structure, to find relevant structure, to retrieve the aligning parts, and to substitute the matching part with something else. Sequence patterns are often described using regular expressions and matched using respective algorithms.

Database Communication :

In this module, independent segments of the database are searched on each processor or node, and results are collated into a single output file and several implementations of database segmentation exist. Database segmentation has also been

implemented in a closed-source commercial product. The database checks for the input to the appropriate Kilo Bytes and if it so i.e. if the pattern matches and it obtains the result.

Front End: Java-J2SE/ .Net

Back End: MySQL/SQL Server

Hardware Requirements: A system with capability of supporting a web server & internet.