Abstract

Smith-Waterman algorithm is one of the most significantly used algorithm and a well known approach to gain information about unknown genes and proteins for biological research. As execution time and accuracy is of great significance as handling large-scale dataset, a more reliable high-throughput and efficient parallelism can be achieved with the adaptation of grid environment. Adapting Smith-Waterman algorithm with the grid environment brings several concerns regarding fault tolerance, variability in resource performance and workload distribution, application availability etc. The work presented here aims at the development of a Dynamic Smith-Waterman algorithm metascheduler that handles all the specifications of job submission on the grid to the end user for local alignment search. Additionally a web based portal using GridSphere portal framework integrated with Globus 4 and Java Commodity Grid
Kit is developed that reduces the complexity to the end users in accessing, managing and manipulating the grid resources and applications. The main contribution towards Dynamic Smith-Waterman algorithm is the reduction of the total job execution time up to 52% with accuracy up to 99.99% and better resource utilization by 40%. In addition, this work can be used as a template for the development of similar applications in future.

References

- Jason Novotny, Michael Russel, and Oliver Werens. GridSphere: An Advanced Portal Framework. In EUROMICRO &apos;04: Proceedings of the 30thEUROMICRO Conference,
Dynamic Smith-Waterman Algorithm: A High-Performance Grid-Enabled Application Integrated with Globus, GridSphere Portal Framework and CoG Workflow for Performing Biological Local Sequence Alignment


Index Terms

Computer Science

Algorithms

Keywords

Grid Globus GridSphere Java Commodity Grid Kit Scheduling

Smith-Waterman algorithm