Biological Computing on the Grid

Protein folding simulation / Multiscale bio-molecular simulation

biogrid

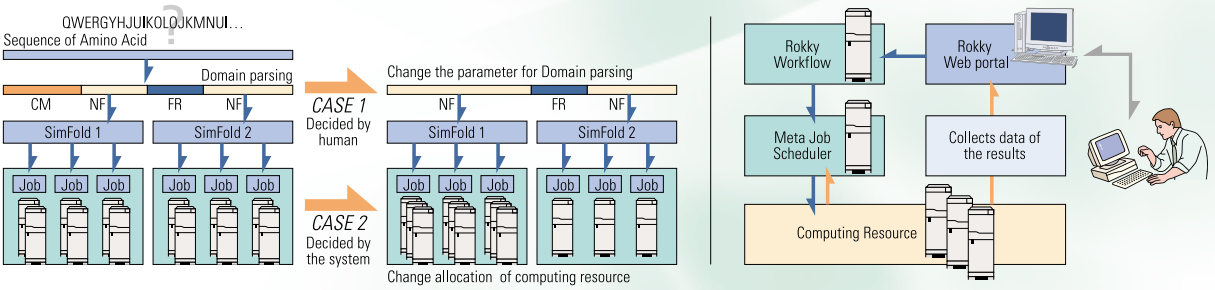
In the post-genome era, revealing protein mechanisms is one of the most important issues for progress in medicine and pharmacology. Biologists demand various types of calculations to solve this issue. We have been studying the requirements of each calculation, and researching approaches to meet the requirements.

HTC : High throughput Computing

Overview

In the scientific simulation, it is necessary to obtain a highly accurate result using limited computer resources. We have devised a Grid System to remove ineffective tasks and prioritize tasks of high importance to the simulation.

In many cases, users can not decide the optimal value of input parameter before executing a task. Therefore the result may not be accurate. If this happens, computer resources have been wasted. When the process report is checked however, the user can find out which parameters worked and which did not. This enables the parameters to be adjusted. This will enable the task to be restarted with a higher level of efficiency.
Additionally during tasks, the computer can choose which tasks require more resources and adjust its allocations accordingly.



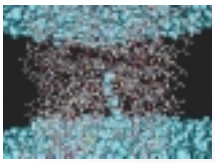
The protein structure forecasts are calculated by dividing the workload into parts. Domain Parsing decides how this is done. The effectiveness of this decision naturally affects the efficiency of the result.

Example cases (please refer to the diagrams above)

CASE 1 shows an example where the user changes the parameters. The user can confirm the status of the task and the result during the simulation at any time. In this example, the parameter of the Domain Parsing has been changed. The execution beginning the task and the change in the parameter are done through the Web portal.
In *CASE 2* the system has automatically changed the computer resource allocations from the status of the task being executed. In the protein structure forecast, the number of task being executed influences the accuracy of the result. In *CASE 2*, the balance of computer resource allocation is adjusted automatically to enable more resources to larger tasks and less resources to smaller tasks. This enables subsequent tasks to be completed at a higher rate of efficiency.
Therefore, the final task in the result will provide the most efficient result.

An application of BioPfuga for hybrid-QM/MM simulation

BioPfuga : Biosimulation Platform United on Grid Architecture



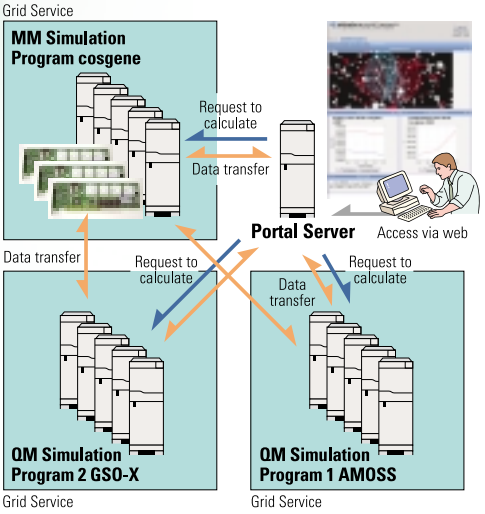
Overview

Today, in order to analyze a highly complicated biological system from multiple perspectives, scientists demand multi-scale simulations. Multi-scale simulations promise to achieve integrated analyses by using different simulation methods simultaneously. These integrated analyses have the potential to introduce a new paradigm into bioinformatics.
We have focused especially on a hybrid-QM/MM simulation, which is one of the multi-scale simulations. The hybrid-QM/MM simulation is a simulation method to simulate bio-molecules based on quantum mechanics (QM) and molecular dynamic mechanics (MM).

Development of an application for hybrid-QM/MM simulation

To carry out this integrated simulation, the use of widely distributed environments or the Grid is effective and efficient. The hybrid-QM/MM simulation consists of different simulation components developed by different organizations. They are independent from each other, and each needs large computational resources. Some simulations also require special devices for calculations. Each component, then, is suitable to run on different environments from the standpoints of performance and cost for preparing running environments.
In order to achieve this integrated simulation, we have developed Grid services or Web services for each simulation based on the OGSA. We have achieved the following two points in this integrated simulation.

◇Development of a standard XML format to describe data of bio-molecular simulation, BMSML
◇Abstraction of data transfer interface by utilizing the inheritance of a Web service interface



These two points have been used to achieve integration of simulations' interfaces while keeping each simulation independent, and then facilitating combinations of simulations. For example, switching simulations, each of which adopts different algorithm, is made easier, as well as mixing the results of these different simulations to achieve higher precision. Consequently, our system has realized complicated combinations of simulation or more sophisticated simulations.

OSAKA UNIVERSITY

http://www.biogrid.jp | Cybermedia Center, Osaka University | 5-1 Mihogaoka Ibaraki Osaka 567-0047 Japan | biogridinfo@ais.cmc.osaka-u.ac.jp