Performance Evaluation of Large-Scale Distributed-Processing System via Grid Datafarm Architecture in the STP field

Kazunori Yamamoto[1]; Ken T. Murata[2]; Daisuke Matsuoka[1]; Eizen Kimura[3]; Osamu Tatebe[4]

[1] Ehime Univ.; [2] CITE, Ehime University; [3] CITE, Ehime Univ.; [4] Computer Science, Univ. of Tsukuba

http://www.infonet.cite.ehime-u.ac.jp/

There are two major research methods for geo-space science; one is computer simulation, and the other is satellite and/or ground-based observation. Both methods have advantages and disadvantages: Computer simulations can provide data in whole time and space in the simulation domain. On the other hand, satellite observation data are hopefully more accurate. Therefore it is effective for confidence and possibility of data analyses to use multi-satellite observation data and to visualize in high time resolution.

It is necessary to solve the following two problems to achieve data-intensive processing for plenty of data files with non-negligible file I/O. (i) More CPU power and larger memory and disk size are required. However, total powers of personal computers are not enough to analyze such amount of data. Super-computers provide a high performance CPU and rich memory area, but they are usually separated from the Internet or connected only for the purpose of programming or data file transfer. (ii) Plenty of data files are distributed among nodes. Users have to know where the data files are located.

To overcome the problems (i) and (ii), we constructed a parallel and distributed data analysis environment based on the Gfarm reference implementation of the Grid Datafarm architecture at Ehime university as shown in figure. The Gfarm shares both computational resources and perform parallel distributed processing. In addition, the Gfarm provides the Gfarm filesystem which behaves as a virtual directory tree among nodes. The Gfarm environment is composed of three parts; a metadata server to manage distributed files information, file system nodes to provide computational resources and a client to throw a job into metadata server and manages data-processing scheduling. Present Gfarm environment is composed of high spec 8 file system nodes.

We examined performance of both sequential processing and parallel processing in the present system. In the experiment, Gfarm performs scanning processing of long durational observation data files. We used 4 types of observation data files. There were various data files from 100KB to 30MB. In the result of experiment, the present method was effective when one data file size is not small enough. As a result of load of swap can be distributed, long durational data can be processed at once.

Next, we performed data processing of computer simulation data. We attempted to shorten the creation time of 3-D object files using Gfarm. Initially, during the time steps from 1 to 150, data files are distributed time step evenly in order of nodes. Secondly, 3-D object files are created by each node. Finally, many created 3-D object files are combined into one animation file. In the result of experiment, it was a bottleneck node in the relation between time step and elapsed time in each node. In this case, there was a node which nothing processed for 136 minutes. It shows this method cannot get load balance effectively. In the ideal, total elapsed time of parallel processing should be one eighth, but the result did not become like it. We paid attention to file distribution and used the method of changing a node for every time step in order of nodes which processing finished. This method was effective when elapsed time differs in every process. Even if users construct hetero environment which is composed of several different machine spec, users can get load balance effectively. Using this method, we attempted to create a 3D animation file whose time resolution is 1 second and the number of time steps is 2,400. To create this 3D animation, it took 1 day even on the Gfarm system. If we create it on one machine, it took for 1 week.

In the present talk, we show latest results and discuss the present Gfarm system extended to global area.