

Grid Computing in Distributed GIS

By Zahid Imran Ahmed

Grid Computing

Some consider this to be the “the third information technology wave” after the Internet and Web, and will be the backbone of the next generation of services and applications that are going to further the research and development of GIS and related areas.

Grid computing allows for the sharing of processing power, enabling the attainment of high performances in computing, management and services. Grid computing, (unlike the conventional supercomputer that does parallel computing by linking multiple processors over a system bus) uses a network of computers to execute a program. The problem of using multiple computers lies in the difficulty of dividing up the tasks among the computers, without having to reference portions of the code being executed on other CPUs.

Parallel processing

Parallel processing is the use of multiple CPU's to execute different sections of a program together. Remote sensing and surveying equipment have been providing vast amounts of spatial information, and how to manage, process or dispose of this data have become major issues in the field of Geographic Information Science (GIS).

To solve these problems there has been much research into the area of parallel processing of GIS information. This involves the utilization of a single computer with multiple processors or multiple computers that are connected over a network working on the same task. There are many different types of distributed computing, two of the most common are clustering and grid processing.

The primary reasons for using parallel computing are:

Saves time.

Solve larger problems.

Provide concurrency (do multiple things at the same time).

Taking advantage of non-local resources - using available computing resources on a wide area network, or even the Internet when local computing resources are scarce.

Cost savings - using multiple "cheap" computing resources instead of paying for time on a supercomputer.

Overcoming memory constraints - single computers have very finite memory resources. For large problems, using the memories of multiple computers may overcome this obstacle.

Limits to serial computing - both physical and practical reasons pose significant constraints to simply building ever faster serial computers.

Limits to miniaturization - processor technology is allowing an increasing number of transistors to be placed on a chip. However, even with molecular or atomic-level components, a limit will be reached on how small components can be.

Economic limitations - it is increasingly expensive to make a single processor faster. Using a larger number of moderately fast commodity processors to achieve the same (or better) performance is less expensive.

The future: during the past 10 years, the trends indicated by ever faster networks, distributed systems, and multi-processor computer architectures (even at the desktop level) clearly show that parallelism is the future of computing.

Distributed GIS

As the development of GIS sciences and technologies go further, increasingly amount of geospatial and non-spatial data are involved in GISs due to more diverse data sources and development of data collection technologies. GIS data tend to be geographically and logically distributed as well as GIS functions and services do. Spatial analysis and Geocomputation are getting more complex and computationally intensive. Sharing and collaboration among geographically dispersed users with various disciplines with various purposes are getting more necessary and common. A dynamic collaborative model - "Middleware" - is required for GIS application.

Computational Grid is introduced as a possible solution for the next generation of GIS. Basically, the Grid computing concept is intended to enable coordinate resource sharing and problem solving in dynamic, multi-organizational virtual organizations by linking computing resources with high-performance networks. Grid computing technology represents a new approach to collaborative computing and problem solving in data intensive and computationally intensive environment and has the chance to satisfy all the requirements of a distributed, high-performance and collaborative GIS. Some methodologies and Grid computing technologies as solutions of requirements and challenges are introduced to enable this distributed, parallel, and high-throughput, collaborative GIS application.

Security

Security issues in such a wide area distributed GIS is critical, which includes authentication and authorization using community policies as well as allowing local control of resource. Grid Security Infrastructure (GSI), combined with GridFTP protocol, makes sure that sharing and transfer of geospatial data and Geoprocessing are secure in the Computational Grid environment.

Conclusion

As the conclusion, Grid computing has the chance to lead GIS into a new “Grid-enabled GIS” age in terms of computing paradigm, resource sharing pattern, online collaboration.