

Dynamic Load Balancing of Scientific Applications on Parallel and Distributed Systems

Large-scale simulation is an important method in scientific and engineering research, as a compliment to theory and experiment, and has been widely used to study complex phenomena in many disciplines. Many large-scale applications are adaptive in that their computational load varies throughout the execution and causes uneven distribution of the workload at run-time. Dynamic load balancing (DLB) of adaptive applications involves in efficiently partitioning of the application and then migrating of excess workload from overloaded processors to underloaded processors during execution.

Different applications have different adaptive characteristics, which may result in the existing DLB schemes/tools not well suited for them. We have been working with scientists from various disciplines with the objective to provide efficient and effective load balancing techniques for their applications. The applications that we are working with include the cosmology application ENZO and the molecular dynamics application GROMACS.

Most existing partition schemes are targeted for homogeneous parallel systems, which are not appropriate for large-scale applications running on heterogeneous distributed systems. Furthermore, the cost entailed by workload migration may consume orders of magnitude more time than the actual partitioning when the excess workload is transferred across geographically distributed machines. In particular, with workload migration, it is critical to take into account that the wide area network (used to connect the geographically distributed sites) performance is dynamic, changing throughout execution, in addition to the resource heterogeneity. To address these problems, we have designed novel data partition and migration techniques that can be applicable to a range of large-scale adaptive applications executed on heterogeneous distributed computing environments.

This work was supported by US National Science Foundation grant NGS-0406328, National Computational Science Alliance with NSF PACI Program, and TeraGrid Resource Allocation.

Members:

- Zhiling Lan
- Yawei Li

Collaborators:

- Valerie Taylor (TAMU)
- X. Sun (IIT)
- Larry Scott (IIT)
- Michael Norman (UCSD)
- Greg Bryan (Columbia Univ.)

Publications:

- Z. Lan, V. Taylor, and Y. Li, "DistDLB: Improving the Efficiency of Cosmology Simulations in Distributed Computing Environments through Hierarchical Load Balancing", *Journal of Parallel and Distributed Computing*, 2006.
- Y. Li and Z. Lan, "A Novel Workload Migration Scheme for Heterogeneous Distributed Computing", *Proc. of 5th IEEE/ACM International Symposium on Cluster Computing and the Grid(CCGrid05)*, Cardiff, UK, 2005.
- Z. Lan, V. Taylor, and G. Bryan, "A Novel Dynamic Load Balancing Scheme for Parallel Systems", *Journal of Parallel and Distributed Computing (JPDC)*, Vol 62/12, pp.1763-1781, 2002.

Contact:

Dr. Zhiling Lan (lan AT iit DOT edu)