Heuristic dynamic load balancing algorithm applied to the fragment molecular orbital method

Yuri Alexeev
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439, USA
yuri@alcf.anl.gov

Prasanna Balaprakash
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439, USA
pbalapra@mcs.anl.gov

ABSTRACT
Load balancing for large-scale systems is an important NP-hard difficulty problem. We propose a heuristic dynamic load-balancing algorithm (HDLB), employing covariance matrix adaptation evolution strategy (CMA-ES), a state-of-the-art heuristic algorithm, as an alternative to the default dynamic load balancing (DLB) and previously developed heuristic static load balancing algorithms (HSLB). The problem of allocating CPU cores to tasks is formulated as an integer nonlinear optimization problem, which is solved by using an optimization solver. On 16,384 cores of Blue Gene/Q, we achieved an excellent performance of HDLB compared to the default load balancer for an execution of the fragment molecular orbital method applied to model protein system quantum-mechanically. HDLB is shown to outperform default load balancing by at least a factor of 2, thus motivating the use of this approach on other coarse-grained applications.

Categories and Subject Descriptors
D.1.3 [**Software**]: Programming techniques – *parallel programming.*

General Terms
Algorithms.

Keywords
Load balancing, non-linear optimization, heuristic algorithms, quantum chemistry.

1. INTRODUCTION
Achieving an even load balance is a key issue in parallel computing, and increasingly so as we enter the exascale supercomputing era in the next decade. By Amdahl's law, the scalable component of the total wall time shrinks as the numbers of processors increases, while the load imbalance, together with the constant sequential component, acts to retard the scalability. Although parallelization of sequential code often requires rewriting the code, adopting an efficient load-balancing scheme can be a simple and effective way to boost scalability and performance.

2. HEURISTIC DYNAMIC LOAD BALANCING ALGORITHM
We adopt Covariance Matrix Adaptation Evolution Strategy (CMA-ES), a state-of-the-art heuristic evolutionary optimization algorithm for dynamic load balancing.

Like any other evolutionary algorithm, CMA-ES performs the following steps: i) Initialize parent population; ii) Generate offspring population by using selection, recombination, mutation operators; iii) Select new parent population by combining the parent and offspring population

CMA-ES uses a distribution for mutation which is generated according to a covariance matrix C. This corresponds to learning a second-order model of the objective function.

Consequently, mutations can adapt to the shape of the objective function landscape and convergence to the optimum can be increased significantly.

Statistics collected over the generations are used to control covariance matrix and other algorithmic parameters.

In order to handle the integer parameters, the decision point is rounded to the nearest integer only during the objective function evaluation.

CMA-ES is used iteratively as follows:

1. Gather load-balancing data from previous iteration

2. Fit a simple linear speedup model for each fragment (when a fragment f runs for s sec with n processes, the product of n and s gives the time for single process)

3. Use the linear speedup models for all fragments as an objective function for CMA-ES

4. Deploy CMA-ES to find the best load-balancing allocation and reallocate the resources

3. RESULTS
The calculations have been done on 1 rack of Blue Gene/Q supercomputer located at Argonne National Laboratory. For load-balancing tests, we chose the protein which consists of 76 fragments (amino acid residues). We created 76 GDDI groups (i.e., number of MPI subgroups) equal to the number of fragments.

On Figure 1 the scaling of our developed HDLB algorithm is shown, and compared against performance of the default dynamic load-balancing (DLB) algorithm and previously developed heuristic load-balancing algorithm (HSLB) [1]. Like HSLB, HDLB on average outperforms default DLB by an order of two. But HDLB is much easier to use, and we plan to implement HDLB directly in the GAMESS.

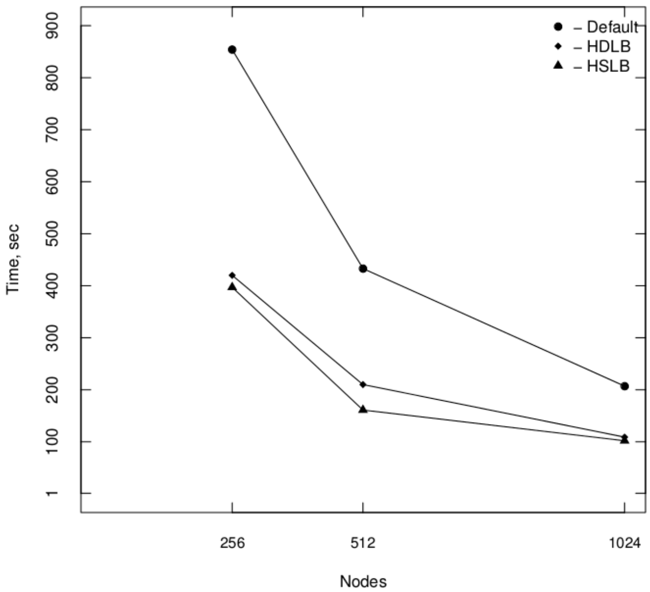

Figure 1. Comparison of scaling HDLB with default dynamic load balancing (DLB) and HSLB.

4. CONCLUSIONS

We have shown that the present HDLB method is twice as fast as the default DLB method and has comparable performance to HSLB. HSLB algorithm provides an optimal solution, but it is hard to use, especially in community codes like GAMESS because MINLP uses a commercial package AMPL. On the other hand, HDLB utilizes an effective and simple algorithm to achieve the same goal. However, there is evidently room for improvement. We plan to continue our work trying different objective functions and also expanding HDLB approach to FMO dimer calculations.

5. ACKNOWLEDGMENTS

This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-ACO2-06CH11357.

6. REFERENCES

[1] "Heuristic static load-balancing algorithm applied to the Fragment Molecular Orbital method", Y. Alexeev, A. Mahajan, S. Leyffer, G. Fletcher, D. G. Fedorov, in Proceedings of the 2012 International Conference for High Performance Computing and Communications (SC 2012), Salt Lake City, UT, USA, (2012).