

Inspector/Executor Load Balancing Algorithms for Block-Sparse Tensor Contractions

David Ozog, Sameer Shende,
Allen Malony
Dept. of Computer and Information Science
University of Oregon
Eugene, Oregon 97403
{ozog, sameer, malony}@uoregon.edu

Jeff R. Hammond, James Dinan,
Pavan Balaji
Mathematics and Computer Science Division
Argonne National Laboratory
Lemont, Illinois
{jhammond, dinan, balaji}@anl.gov

ABSTRACT

Developing effective yet scalable load-balancing methods for irregular computations is critical to the successful application of simulations in a variety of disciplines at petascale and beyond. This poster explores a set of static and dynamic scheduling algorithms for block-sparse tensor contractions within the NWChem computational chemistry code for different degrees of sparsity (and therefore load imbalance). In this particular application, a relatively large amount of task information can be obtained at minimal cost, which enables the use of static partitioning techniques that take the entire task list as input. However, fully static partitioning is incapable of dealing with dynamic variation of task costs, such as from transient network contention or operating system noise, so we also consider hybrid schemes that utilize dynamic scheduling within subgroups. These two schemes, which have not been previously implemented in NWChem or its proxies (i.e. quantum chemistry mini-apps) are compared to the original centralized dynamic load-balancing algorithm as well as improved centralized scheme. In all cases, we separate the scheduling of tasks from the execution of tasks into an inspector phase and an executor phase. The impact of these methods upon the application is substantial on a large InfiniBand cluster: execution time is reduced by as much as 50% at scale. The technique is applicable to any scientific application requiring load balance where performance models or estimations of kernel execution times are available.

Categories and Subject Descriptors

J.2 [Computer Applications]: Physical Sciences and Engineering

General Terms

Algorithms, Performance

Keywords

Dynamic Load Balancing, Static Partitioning, Tensor Contractions, Quantum Chemistry, Global Arrays