

Exploring performance improvement for Java-based scientific simulations that use the Swarm Toolkit

Xiaorong Xiang Gregory Madey

Computer Science & Engineering
University of Notre Dame

Abstract

There is growing interest in using the Java programming language for scientific and engineering applications. This is because Java offers several features, which other traditional languages (C, C++, Fortran) lack, including portability, "garbage collection" mechanism, build-in threads, and the RMI mechanism. However, the historic poor performance of Java stops it from being widely used in scientific applications. Although research and development on Java resulting in JIT compilers, JVM improvement and high performance compilers, have been done much for runtime environment optimization and significantly speeded up the Java programs, we still believe that creating hand-optimized Java source code and optimizing code for speed, reliability, scalability and maintainability are also crucial during program development stage.

Natural organic matter (NOM), a mixture of organic molecules with different types of structure and composition, micro-organisms and their environment form a complex system. The NOM simulator is an agent-based stochastic model for simulating the behaviors of molecules over time. The simulator is built using Java programming language and the Swarm agent-based modeling library. We analyze the NOM simulation model from several aspects: runtime optimization, database access, objects usage, parallel and distributed computing. The NOM simulation model possesses most of characteristics which general scientific applications have. These techniques and analysis approaches can be generally used in other scientific applications. We expect that our experiences can help other developers using Java/Swarm to find a suitable way of achieving higher performance for their applications.