Computer Physics Communications 204 (2016) 132–140

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Scalability Test of multiscale fluid–platelet model for three top supercomputers

Peng Zhang [a,*], Na Zhang [b], Chao Gao [a], Li Zhang [b], Yuxiang Gao [c], Yuefan Deng [b], Danny Bluestein [a,*]

[a] Biomedical Engineering Department, Stony Brook University, Stony Brook, NY 11794-3600, United States

[b] Applied Mathematics Department, Stony Brook University, Stony Brook, NY 11794-8151, United States

[c] Cluster Solution Department, Cray Inc., San Jose, CA 95112, United States

ARTICLE INFO

Article history:
Received 2 September 2015
Received in revised form
18 February 2016
Accepted 30 March 2016
Available online 8 April 2016

Keywords:
Performance analysis
Heterogeneous multicore and multi-GPU architecture
Multiscale simulation
Computational bioengineering

ABSTRACT

We have tested the scalability of three supercomputers: the Tianhe-2, Stampede and CS-Storm with multiscale fluid–platelet simulations, in which a highly-resolved and efficient numerical model for nanoscale biophysics of platelets in microscale viscous biofluids is considered. Three experiments involving varying problem sizes were performed: Exp-S: 680,718-particle single-platelet; Exp-M: 2,722,872-particle 4-platelet; and Exp-L: 10,891,488-particle 16-platelet. Our implementations of multiple time-stepping (MTS) algorithm improved the performance of single time-stepping (STS) in all experiments. Using MTS, our model achieved the following simulation rates: 12.5, 25.0, 35.5 μs/day for Exp-S and 9.09, 6.25, 14.29 μs/day for Exp-M on Tianhe-2, CS-Storm 16-K80 and Stampede K20. The best rate for Exp-L was 6.25 μs/day for Stampede. Utilizing current advanced HPC resources, the simulation rates achieved by our algorithms bring within reach performing complex multiscale simulations for solving vexing problems at the interface of biology and engineering, such as thrombosis in blood flow which combines millisecond-scale hematology with microscale blood flow at resolutions of micro-to-nanoscale cellular components of platelets. This study of testing the performance characteristics of supercomputers with advanced computational algorithms that offer optimal trade-off to achieve enhanced computational performance serves to demonstrate that such simulations are feasible with currently available HPC resources.

Published by Elsevier B.V.

1. Introduction

Numerical simulations revolutionize many fields of engineering and science [1,2] by extending traditional theoretical studies and laboratory experiments to explore multiscale phenomena that are hardly observable or measurable in laboratory settings. Such simulations also require overcoming computational challenges including resolving the diverse spatial–temporal scales [3–5] for understanding multi-component biological and behavioral systems [3–5]. For instance, cardiovascular diseases account for nearly 30% of all deaths globally and 35% of all US deaths annually. Whether due to acute thrombosis associated with myocardial infarction, or progressive intermittent atherothrombotic events, significant ventricular dysfunction may result, leading to hear failure. Presently over 5.5 million patients suffer from heart failure in US and their number is expected to grow by 50%. Of these patients, a significant proportion will become candidates for mechanical circulatory support and prosthetic cardiovascular devices, also burdened with thromboembolic risk and complications. The coagulation cascade of blood may be initiated by flow-induced platelet activation, which prompts clot formation in prosthetic cardiovascular devices and in arterial disease processes. Upon activation, platelets undergo complex biophysical and morphological changes. Activated platelets polymerize fibrinogen into a fibrin network that enmeshes red blood cells. Continuum methods fail to capture the small-scale molecular mechanisms such as filopodia formation upon platelet activation. Utilizing molecular dynamics (MD) that can capture and model the molecular mechanisms is computationally prohibitive for large-scale hematologic problems [5,6]. Innovative numerical approaches are essential for elucidating such vexing problems and enhance the linkage to applications for solving complex problems at the interface of biology

* Correspondence to: Biofluids Research Group, Biomedical Engineer Department, Stony Brook University, NY 11794, United States.
E-mail addresses: peng.zhang@stonybrook.edu (P. Zhang), danny.bluestein@stonybrook.edu (D. Bluestein).

http://dx.doi.org/10.1016/j.cpc.2016.03.019
0010-4655/Published by Elsevier B.V.

and engineering. We propose a multiscale numerical approach for modeling the multiscale fluid–platelet phenomena. The proposed approach utilizes a multiscale model where dissipative particle dynamics is coupled with molecular dynamics to describe flow induced mechanotransduction processes and biochemical events spanning the vast range of spatial and temporal scales characterizing blood flow-induced clotting and thrombosis [6–9].

However, the challenge of multiscale modeling cannot be simply resolved by leveraging on the raw computing speeds of hardware systems. The multiscale nature of such problems mandates developing efficient computational algorithms Hence, we develop a multiple time-stepping (MTS) algorithm in which time integration is carried out at multiple temporal scales [8] and we implement this algorithm on multi-CPU and multi-GPU supercomputers. Our MTS algorithm helps to bridge the temporal gap between macroscale flow regime and cellular-scale events. As a result, it avoids a massive number of redundant computations that are inherent in single time-stepping (STS) algorithms, thus significantly boosting the computing efficiency.

In terms of the scalability tests, we build a problem-specific benchmark for three benchmark systems. The smallest system contains 680,718 particles, forming a single platelet that flows in a viscous flow. In this system, the platelet and flow models contain 140,015 (21%) and 540,703 (79%) particles, respectively. This system is typically used for parameterization and observation of morphological changes for a single human platelet. The largest system contains 10,891,488 particles and it includes 16 platelets. This system helps to study the initial stage of clot processes involving multiple platelets. The middle system contains 2,722,872 particles forming 4 platelets. This system is suitable for validating the inter-platelet and platelet–vessel interactions. From a practical point of view, our results provide the references for evaluating the computational capabilities for given computers. These benchmarks are carried out on two main categories of supercomputer architectures: homogeneous multiprocessor and heterogeneous multi-accelerator architectures. This selection allows us to establish the relationship between diverse supercomputer architectures and diverse multiscale simulations.

In terms of engineering efforts, we implement our novel multiscale model and 4-level multiple time-stepping (MTS) algorithm on three supercomputers of different architectures and we further implement this MTS algorithm on GPUs. These implementations enabled us to: provide for the first time computational insights for solving multiscale simulations; and demonstrate the computational capabilities for simulating *millisecond*-scale multiscale phenomena of multi-component biological systems. The substantial speedup resulting from aggregated efforts of physical models, mathematical algorithms and sophisticated implementation on supercomputers of different architectures offer the community of computational science and biomedical engineering new insights and enhanced ability for incorporating and advancing multiscale modeling in their research. Our work is the first examination of the performance of multiscale platelet-mediated simulations on supercomputers. It clearly demonstrates the feasibility and affordability for conducting millisecond-order micro-to-nanoscale simulations of multiple platelets in viscous blood flows on top supercomputers, an essential step for numerical studies of the thrombogenicity in clinical problems. The integration of multiple model and MTS algorithm is a cost-effective approach for various classes of specific problems in biomedical engineering and biomedical sciences.

The computational research presented in this paper is organized as follows: a multiscale model and multiple time-stepping (MTS) algorithm is presented in Section 2. The key features of three supercomputers are described in Section 3. In Section 4, the experiment setup is presented, followed by results. In Section 4.3, we characterize and analyze the performance of individual supercomputers, followed by conclusions in Section 5.

2. Multiscale simulations

The multiscale model of a platelet is introduced by examining the anatomy of the platelet. The numerical algorithm is designed to model the platelet with sufficient spatio-temporal resolution using minimal computing resources that would typically require far larger ones. The key element of the algorithm is the multiple time-stepping integration of the equation of motion. In this section, we first describe the multiscale model where multiple spatial–temporal scales are formulated and then employ the multiple time-stepping (MTS) algorithm as an efficient numeric solver for this model.

2.1. Multiscale model

By extending our previous efforts for modeling platelets under viscous shear flow conditions at multiple spatio-temporal scales [6–9], we develop a multiscale benchmark model for assessing the performances of three supercomputers for understanding multiscale phenomena of complex biological systems. In this model, we cover two spatio-temporal scales: (i) the microscale flow regime using dissipative particle dynamics (DPD) to describe the bulk transfer of viscous blood flow [10]; and, (ii) the nanoscale platelet model using coarse-grained molecular dynamics (CGMD) to describe the cellular scale structural details of a platelet such as membranous morphology, cytoplasmic biorheology, cytoskeletal filaments and the flow-mediated cellular mechanotransduction of hemodynamic stresses across the platelet surface and through the cytoskeleton [6]. Both DPD and CGMD are coarse-grained particle-based methods. Specifically DPD/CGMD is the coarse-grained stochastic/molecular dynamics, respectively. In the formulation, DPD employs a special stochastic character that depends on the momentum to increase the scales of viscous flows it can formulate. CGMD employs a conservative force field to formulate the cellular-scale intra-platelet interactions. The interface of these two formulations is described by a hybrid force field; the local thermodynamics and exchange of momentum is governed by the DPD stochastic term while the incompressibility of the platelets, under the stress of the surrounding flow is governed by the CGMD conservative term [6]. Fig. 1 is a schematic description of the multiple spatio-temporal scales in which different force fields capture the characteristics of the different scales. A wall-driven Couette benchmark flow is introduced to induce the characteristic flipping of the deformable platelets in such shear flows.

This model may be utilized to study similar phenomena of flowing blood cells such as erythrocytes and leukocytes that require understanding, a priori, the scalability and time-to-solution for such problems. Fig. 2 illustrates the hierarchy of multiple spatial–temporal scales correlated with biological phenomena at multiple resolutions.

2.2. Multiple time stepping algorithm

The spatial interface between DPD and CGMD is the first step for performing such multiscale simulation and the corresponding time integration, multiple time stepping (MTS) algorithm, solving the DPD and CGMD at microseconds and nanoseconds temporal scales [8] is the following step. Far more sophisticated than single time stepping algorithms [11–13], MTS algorithms handle time integration of step sizes differing by 3–4 orders of magnitude, as those between DPD and CGMD time scales. Following [8], we use a four-step scheme model in which the fluid advances at the largest step size, the fluid–platelet interface at a middle step size and the nonbonded and bonded force fields within platelets at the two smallest step sizes. The relationship between accuracy and computational load, i.e., the MTS parameters, was studied in [8].