Parallel Computing 49 (2015) 1–12

Contents lists available at ScienceDirect

Parallel Computing

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

Improved strong scaling of a spectral/finite difference gyrokinetic code for multi-scale plasma turbulence

Shinya Maeyama[a,b,*], Tomohiko Watanabe[b], Yasuhiro Idomura[a], Motoki Nakata[c], Masanori Nunami[d], Akihiro Ishizawa[d]

[a] Japan Atomic Energy Agency, 5-1-5 Kashiwanoha, Kashiwa 277-8587, Japan
[b] Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
[c] Japan Atomic Energy Agency, 2-166 Omotedate, Obuchi, Rokkasho, Kamikita-gun 039-3212, Japan
[d] National Institute for Fusion Science/The Graduate University for advanced Studies, 322-6 Oroshi-cho, Toki 509-5292, Japan

ARTICLE INFO

Article history:
Received 16 May 2014
Revised 20 March 2015
Accepted 1 June 2015
Available online 9 June 2015

Keywords:
Gyrokinetics
Computational fluid dynamics application
Parallel scalability
Spectral method
Finite difference

ABSTRACT

Optimization techniques of a plasma turbulence simulation code GKV for improved strong scaling are presented. This work is motivated by multi-scale plasma turbulence extending over multiple spatio-temporal scales of electrons and ions, whose simulations based on the gyrokinetic theory require huge calculations of five-dimensional (5D) computational fluid dynamics by means of spectral and finite difference methods. First, we present the multi-layer domain decomposition of the multi-dimensional and multi-species problem, and segmented MPI-process mapping on 3D torus interconnects, which fully utilizes the bi-section bandwidth for data transpose and reduces the conflicts of simultaneous point-to-point communications. These techniques reduce the inter-node communication cost drastically. Second, pipelined computation-communication overlaps are implemented by using the OpenMP/MPI hybrid parallelization, which effectively mask the communication cost. Careful regulations of the pipeline length and of the thread granularity respectively suppress latencies of MPI, load imbalance and scheduling overheads of OpenMP. Thanks to the above optimizations, GKV achieves excellent strong scaling up to ∼600k cores with high computational performance 782.4 TFlops (8.29% of the theoretical peak) and high effective parallelization rate ∼99.99994% on K, which demonstrates its applicability and efficiency toward a million of cores. The optimized code realizes multi-scale plasma turbulence simulations covering electron and ion scales, and reveals cross-scale interactions of electron- and ion-scale turbulence.

1. Introduction

Nuclear fusion energy is expected to be a future energy source because of huge resources and manageable environmental impacts. The most advanced approach to the fusion energy is based on magnetic confinement of plasma in toroidal devices, where turbulent electromagnetic fluctuations strongly influence the confinement properties. Therefore, plasma turbulent transport is one of the central issues in the magnetic fusion research. The plasma turbulence is inherently a multi-scale phenomenon involving electron and ion scales, which are separated by a factor of a square root of the mass ratio when their temperatures are the same (typically, that is $\sqrt{m_\mathrm{i}/m_\mathrm{e}} \sim 43$). Additionally, since low collisionality in high temperature plasma makes kinetic

* Corresponding author at: Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan. Tel.: +81 52 789 3930.
E-mail address: smaeyama@p.phys.nagoya-u.ac.jp (S. Maeyama).

http://dx.doi.org/10.1016/j.parco.2015.06.001
0167-8191/

effects essentially important, plasma turbulence is necessary to be described by means of particle distribution functions defined on the phase space (the configuration and the particle-velocity spaces). Such multi-dimensional simulations demand huge computational costs, and the multi-scale analysis requires further expensive computations.

Fundamental understandings of the plasma turbulence are established based on analyses of the locally homogeneous turbulence by using spectral methods [1], as is the case in turbulence in neutral fluids. The spectral methods have excellent numerical accuracy, such as exponential convergence and no numerical dissipation error. The Fourier expansion is compatible to theoretical analysis of energy transfer via triad wave interactions, and its computational cost is saved by applying fast Fourier transform (FFT) algorithms. Parallel multi-dimensional FFTs, however, are often implemented with transpose split method [2] and introduce global data transpose communications which often degrade parallel scalability. Regarding applications based on the spectral method, parallelization on ten thousands of cores (a few thousands of computation nodes) was realized [3], but it was considered that the spectral method is hardly scalable as the number of cores is increased up to hundreds of thousands.

In this paper, we present the parallelization of spectral/finite difference calculations in gyrokinetic plasma turbulence simulation code GKV [4,5] for improved strong scaling up to ∼600 k cores. To bring out the application performance under the parallelization on hundreds of thousands of cores, deep understanding of the target problem and its tailored optimizations on hierarchical parallel architectures are necessary. We design the multi-layer domain decomposition and topology-aware MPI-process mapping to minimize inter-node communications, and communication overlaps with combining various physical computations to mask the communication cost effectively. The optimized code, of which performance is strongly accelerated, is applied to the analysis of the multi-scale plasma turbulence, revealing the importance of cross-scale interactions between electron- and ion-scale turbulence with the real hydrogen-to-electron mass ratio.

This paper is organized as follows. Section 2 describes the basic simulation model, as well as the hybrid parallelization which consists of four-dimensional (4D) and multi-species domain decomposition with MPI and shared memory parallelization with OpenMP. Section 3 explains the segmented MPI-process mapping and pipelined computation-communication overlaps which are designed for the optimization of the multi-dimensional and multi-species problem to the hierarchical architecture connected by 3D torus interconnects. These sophisticated parallelization techniques improve strong scaling of the code and enhance computational performance, as shown in Section 4. In Section 5, multi-scale turbulence simulations are carried out, which demonstrates the practical utility of the developed code as an application to solve physically-motivated problems demanding huge computations. Finally, we summarize the developed techniques and give some concluding remarks in Section 6.

2. The GKV code

Turbulence in magnetic fusion plasma is characterized by anisotropy due to the strong confinement magnetic field. Perpendicular and parallel spatial scales of plasma turbulence are respectively in the order of gyroradii of charged particles ρ and the major radius of the toroidal device R (as an example $\rho/R \sim 10^{-3}$ for ions in ITER), where, and hereafter, "parallel" and "perpendicular" mean the directions oriented to the confinement magnetic field unless otherwise specified. To investigate the plasma turbulence, the gyrokinetic theory [6] has been developed, where rapid gyrations are decoupled from the slow dynamics of turbulent fluctuations while resolving spatial scales of gyroradii. The GKV code is based on the so-called δf gyrokinetics [7] and solves time evolution of perturbed distribution functions $\tilde{f}_\mathrm{s}$ of ions and electrons (the subscript $\mathrm{s} = \mathrm{i}, \mathrm{e}$), the electrostatic potential $\tilde{\phi}$ and the parallel component of the vector potential $\tilde{A}_\parallel$. The governing equations are the gyrokinetic Vlasov–Poisson–Ampère equations

$$\frac{\partial \tilde{f}_\mathrm{s}}{\partial t} + \left(v_\parallel \frac{\boldsymbol{B} + \tilde{\boldsymbol{B}}_\perp}{B} + \boldsymbol{v}_\mathrm{sG} + \boldsymbol{v}_\mathrm{sC} + \tilde{\boldsymbol{v}}_\mathrm{E} \right) \cdot \nabla \tilde{f}_\mathrm{s} + \frac{dv_\parallel}{dt} \frac{\partial \tilde{f}_\mathrm{s}}{\partial v_\parallel} = S_\mathrm{s} + C_\mathrm{s}, \tag{1}$$

$$\nabla_\perp^2 \tilde{\phi} = -\frac{1}{\varepsilon_0} \sum_\mathrm{s} e_\mathrm{s} (\tilde{n}_\mathrm{s} + \tilde{n}_\mathrm{s,pol}), \tag{2}$$

$$\nabla_\perp^2 \tilde{A}_\parallel = -\mu_0 \sum_\mathrm{s} e_\mathrm{s} \tilde{u}_{\parallel \mathrm{s}}, \tag{3}$$

where B, $\boldsymbol{B}$, $\tilde{\boldsymbol{B}}_\perp$, $v_\parallel$, $\boldsymbol{v}_\mathrm{sG}$, $\boldsymbol{v}_\mathrm{sC}$, $\tilde{\boldsymbol{v}}_\mathrm{E}$, S_s, C_s, ε_0 and μ_0 denote the strength and vector of the confinement magnetic field, the magnetic perturbation, the parallel, grad-B, curvature and $\boldsymbol{E} \times \boldsymbol{B}$ drift velocities, the contributions from the equilibrium distribution, the model collision operator, the vacuum permittivity and permeability, respectively. $dv_\parallel/dt$ represents the parallel acceleration, where the parallel nonlinearity is neglected [8]. The gyrocenter density $\tilde{n}_\mathrm{s}$ and the parallel flow velocity $\tilde{u}_{\parallel \mathrm{s}}$ are given by the velocity integration of the perturbed distribution function, while the polarization density $\tilde{n}_\mathrm{s,pol}$ originates from the deviation of the particle and gyrocenter positions and is proportional to the electrostatic potential perturbation $\tilde{\phi}$. For more detailed physical descriptions, see [4,5].

From numerical viewpoints, the code performs computational fluid dynamics (CFD) calculations on the 5D phase space (x, y, z, $v_\parallel$, μ) for each species (s), where the perpendicular directions x and y and the parallel direction z construct the magnetic field-aligned spatial coordinates [9], and the parallel velocity $v_\parallel$ and the magnetic moment μ (which corresponds to the perpendicular velocity) are employed as the velocity space coordinates. By assuming the locality and statistic homogeneities in the perpendicular directions x and y, the perturbed quantities are expanded in terms of the Fourier basis: $\tilde{f}_\mathrm{s}(x, y, z, v_\parallel, \mu) = \sum_{\boldsymbol{k}_\perp} \hat{f}_{\mathrm{s}\boldsymbol{k}_\perp}(z, v_\parallel, \mu) e^{i(k_x x + k_y y)}$, $\tilde{\phi}(x, y, z) = \sum_{\boldsymbol{k}_\perp} \hat{\phi}_{\boldsymbol{k}_\perp}(z) e^{i(k_x x + k_y y)}$ and $\tilde{A}_\parallel(x, y, z) = \sum_{\boldsymbol{k}_\perp} \hat{A}_{\parallel \boldsymbol{k}_\perp}(z) e^{i(k_x x + k_y y)}$. Then, the Eqs. (1)–(3) are