

MODYLAS: A Highly Parallelized General-Purpose Molecular Dynamics Simulation Program for Large-Scale Systems with Long-Range Forces Calculated by Fast Multipole Method (FMM) and Highly Scalable Fine-Grained New Parallel Processing Algorithms

Yoshimichi Andoh,[†] Noriyuki Yoshii,[†] Kazushi Fujimoto,[†] Keisuke Mizutani,[†] Hidekazu Kojima,[†] Atsushi Yamada,[†] Susumu Okazaki,[†] Kazutomo Kawaguchi,^{*,‡} Hidemi Nagao,[‡] Kensuke Iwahashi,[§] Fumiyasu Mizutani,[§] Kazuo Minami,^{||} Shin-ichi Ichikawa,[⊥] Hidemi Komatsu,[⊥] Shigeru Ishizuki,[⊥] Yasuhiro Takeda,[⊥] and Masao Fukushima[⊥]

[†]Department of Applied Chemistry, Nagoya University, Nagoya 464-8603, Japan

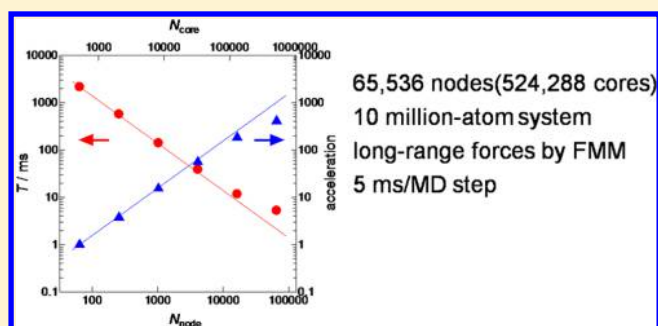
[‡]Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa 920-1192, Japan

[§]Institute for Molecular Science, Okazaki 444-8585, Japan

^{||}Advanced Institute for Computational Science, Riken, Kobe 650-0047, Japan

[⊥]Computational Science and Engineering Solution Division, Fujitsu Limited, Chiba 261-8588, Japan

ABSTRACT: Our new molecular dynamics (MD) simulation program, MODYLAS, is a general-purpose program appropriate for very large physical, chemical, and biological systems. It is equipped with most standard MD techniques. Long-range forces are evaluated rigorously by the fast multipole method (FMM) without using the fast Fourier transform (FFT). Several new methods have also been developed for extremely fine-grained parallelism of the MD calculation. The virtually buffering-free methods for communications and arithmetic operations, the minimal communication latency algorithm, and the parallel bucket-relay communication algorithm for the upper-level multipole moments in the FMM realize excellent scalability. The methods for blockwise arithmetic operations avoid data reload, attaining very small cache miss rates. Benchmark tests for MODYLAS using 65 536 nodes of the K-computer showed that the overall calculation time per MD step including communications is as short as about 5 ms for a 10 million-atom system; that is, 35 ns of simulation time can be computed per day. The program enables investigations of large-scale real systems such as viruses, liposomes, assemblies of proteins and micelles, and polymers.



1. INTRODUCTION

In recent decades, molecular dynamics (MD) simulation has occupied an essential position in the biosciences and material sciences. The development of the method has three trends, which are closely related to the architecture of computers. The first is the extension of the target system to very large systems. It enables investigations of 10–10² nm-scale structures formed by molecules. Typical systems are viruses, cells, and polymers. A massively parallel supercomputer with three-dimensional (3D) torus communication architecture^{1,2} is very powerful for this kind of calculation. The second trend is the long-time dynamics, where the calculation time for one MD step is reduced to the order of microseconds. This enables simulations of millisecond dynamics of, for example, proteins such as protein folding and slow thermal fluctuations. Specific machines such as MD-GRAPE^{3,4} and ANTON⁵ lead this trend. The third

trend is to obtain large statistics. Low-price PC (personal computer) clusters are suitable for this.

The most exciting of the three trends is, of course, the first one. In the present study, we concentrate on the extension of MD calculations to very large systems using massively parallel supercomputers. We believe that this extension can open a new frontier in the science of molecular systems.

However, there are several important conditions to be satisfied in large-scale MD simulations, without which the simulations can be misleading. The first condition is the rigorous evaluation of long-range forces. For example, poliovirus, one of the targets of the present study, has a distribution of negative electric charges on the inner surface of the spherical capsid, which has a diameter of about 30 nm. A receptor

Received: March 14, 2013

Published: June 11, 2013

approaching the virus feels electrostatic forces from the whole body of the virus. If only the short-range forces within, for example, 1.2 nm are taken into account in the simulation, it causes serious errors. The detailed structure of the capsid is not involved in the uniform force correction such as that provided by the reaction field method.

The second condition is the periodic boundary condition. For example, a virus is not in a vacuum but in bulk water. We must therefore impose a periodic boundary condition on the system to remove a surface effect. The calculation describing the system as a droplet in a vacuum is unreliable because very strong surface tension is inevitably included. This is different from astronomical simulations, where the system is described as isolated.

The third condition is that divergent chemical and biological simulations require a general-purpose program that may be applied to various molecular systems ranging from small molecules to polymers. In particular, arbitrary atom–atom distance constraints combined with temperature and pressure control are necessary. Application of various force fields is also desirable.

The fourth condition is the free energy calculation. In general, the time constant of the dynamics of interest found in large molecular systems is much longer than that traceable by the simulation. In this case, the dynamics is often discussed thermodynamically along the path assumed for it. Then, free energy plays a principal role.

Thus, an MD program for large-scale systems should be equipped with (1) rigorous calculation of long-range forces; (2) periodic boundary conditions; (3) applicability to arbitrary molecular systems; and (4) free energy calculations.

In highly parallelized calculations, the first and second requirements—the rigorous evaluation of long-range forces for systems with periodic boundary conditions—are the central problems to be solved. In the conventional MD calculations, the particle-mesh Ewald (PME) method⁶ has been used for this. However, this method requires fast Fourier transform (FFT) calculations, which have been a severe hindrance to be removed for massively parallel supercomputers. The parallelization efficiency for the FFT calculation becomes poor when the parallelism exceeds a certain degree. This is why potential cutoff approximation or uniform force correction based on the reaction field theory has been used. However, these approximations suffer from large errors.

In the present study, the long-range forces acting over the whole system with periodic boundary conditions are evaluated by the fast multipole method (FMM) combined with the Ewald method for the MD unit cell multipoles.^{7,8} The FMM is free from FFT and only requires communication between near neighbors at each level, which is suited to very high parallelism. Because the communication required for the Ewald method for the cell multipoles is very small compared with the PME method, the calculation time is negligible. The accuracy of the calculation can be controlled by the order of spherical harmonic expansion such that we may choose it as we like, depending on the science.

Performance of the program has been tested on the K-computer,⁹ and a very high scalability, and thus a very high performance, has been achieved. The K-computer at RIKEN Advanced Institute for Computational Science, Japan, is one of the foremost supercomputer systems. With 82 944 compute nodes of eight-core processors, it is in the highest class of the powerful supercomputers in the world. One important aspect

of the K-computer is its interconnect system: the virtual 3D torus network (Tofu)⁹ mapped on a six-dimensional network connects the compute nodes of this massively parallel computer.

The theoretical peak performance of the K-computer is 10.61 Pflops. Only some researchers can use it at present. However, supercomputers offering 1 Pflops are now becoming popular, and 10 Pflops-class supercomputers will be readily available in five years, and many researchers will be able to occupy a large number of nodes. In this sense, highly parallelized programs are important. Of course, MODYLAS may be used even in PC clusters, although the performance depends on the degree of parallelization.

Given the requirement of the molecular dynamics simulation, the main challenge for massively parallel processing is in the extremely fine-grained parallel processing. For the new network architecture and the required granularity, new communication methods as well as new data structures are necessary to make the fine-grained parallel processing efficient. Further, a new parallel strategy for the FMM, which tends to require global operations on the upper levels for both communications and arithmetic operations, should be introduced to extract the full potential of the new Tofu interconnect.

The new program designs adopted in the present study are as follows:

- (1) a new data structure to realize virtually buffering-free operation for both communications and arithmetic operations
- (2) a new communication strategy for localized gather operations with minimal communication latency and minimal communication data volume
- (3) a new loop-ordering strategy for greatly enhanced data access locality for arithmetic operations

2. MODYLAS

A highly parallelized general-purpose MD program, MODYLAS (MOlecular DYnamics simulation software for LARGE Systems), written in FORTRAN90, is equipped with most MD techniques. The long-range forces in the periodic boundary condition may be evaluated rigorously by the FMM combined with the Ewald method for the MD unit cell multipoles.^{7,8} MODYLAS is also equipped with PME for small systems, although this is not discussed here. The temperature and pressure may be controlled by the Nosé–Hoover chain^{10,11} and the Andersen method,¹² respectively, generating NVE, NVT, and NPT ensembles. Atom–atom distance constraints may be set arbitrarily by SHAKE/RATTLE/ROLL.^{13–15} The equations of motion are solved by RESPA, where multiple time step calculations may be performed.¹⁶ The program can deal with various force fields such as CHARM22 with CMAP,¹⁷ CHARM36 with CMAP,¹⁸ AMBER,¹⁹ and OPLSAA.²⁰ Free-energy calculation is also available, based on the thermodynamic integration method.

The cubic MD unit cell is divided by 2^n for each side, giving 8^n cubic subcells. Once the number of MPI process has been determined, an equally divided subcell block is assigned to each MPI process. Details of the new methods and algorithms implemented in MODYLAS are presented below.



Figure 1. Hierarchical decomposition of the MD unit cell in the FMM.

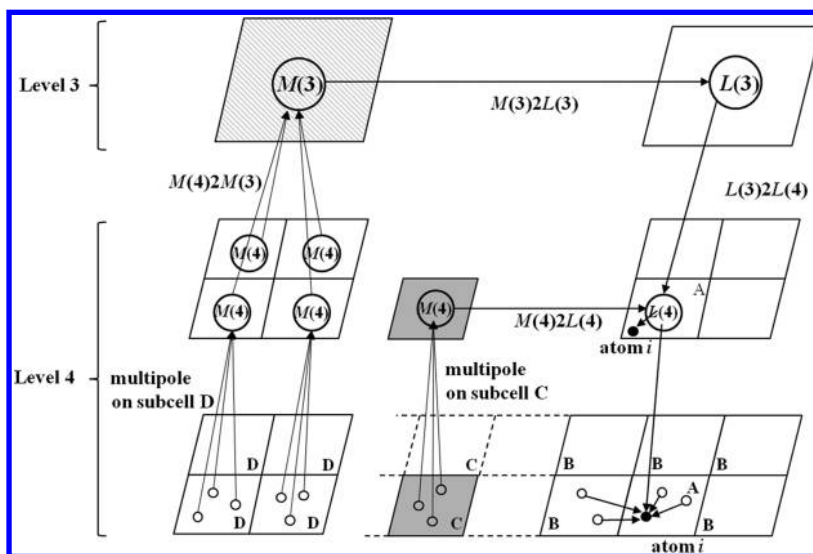


Figure 2. Interaction calculation for the i th atom in subcell A in the FMM.

3. LONG-RANGE FORCES IN THE PERIODIC BOUNDARY CONDITION

3.1. FMM. In this section, we describe how we treat long-range interactions using FMM. Figure 1 shows the two-dimensional schematic view of hierarchical decomposition of the MD unit cell into subcells. Subcells A, B, and C are the smallest. We now evaluate the interactions of an atom i in subcell A. Beyond subcell C, a larger subcell (= supercell) D is formed from eight neighboring smallest subcells in three dimensions (four neighboring subcells in two dimensions in the figure). Furthermore, supercell E is built up from neighboring supercells D. As the distance from the subcell A increases, so does the coarse-graining of the subcells and supercells.

Interactions between an atom i in subcell A and atoms in the neighboring subcell B are evaluated by a direct-pair interaction calculation, as shown in Figure 2. With regard to the subcells C and supercells D, charges in each cell are approximated by a series of multipoles at the center of each cell. The interactions between the atom in subcell A and those in subcell C (Figure 3a) can be obtained by the multipole expansion by the spherical harmonics Y_n^m as

$$V = \frac{1}{4\pi\epsilon_0} \frac{Q}{r} \sum_n \sum_{m=-n}^n M_n^m \frac{Y_n^m(\theta, \varphi)}{r^n} \quad (1)$$

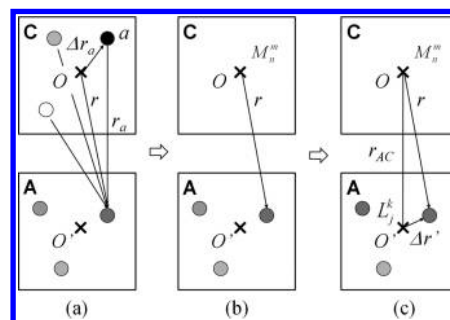


Figure 3. Multipole expansion and local expansion.

where θ and φ are the zenith and azimuth angles, respectively, of the target atom in subcell A with charge Q . The n th multipole moment M_n^m in Figure 3a and b is given by

$$M_n^m = \sum_{a=1} Q_a Y_n^{-m}(\alpha_a, \beta_a) \Delta r_a^n \quad (2)$$

where α_a and β_a are the zenith and azimuth angles, respectively, of the a th atom with charge Q_a located at Δr_a from the origin O in subcell C.

In the FMM, the distance r between the target atom and the origin O is expressed by the local expansion of the spherical

harmonics at the center of subcell A (Figure 3c). The distance r_{AC} between the centers of subcells A and C is independent of the configuration of atoms, so the distance r_{AC} can be included in the local expansion coefficient. Then, the interaction is presented by

$$V = \frac{Q}{4\pi\epsilon_0} \sum_{j=0}^{\infty} \sum_{k=-j}^j L_j^k Y_j^k(\theta', \varphi') \Delta r'^j \quad (3)$$

where L_j^k is the j th local expansion coefficient, which contains the contribution of the multipole of the subcell C. θ' and φ' are the zenith and azimuth angles, respectively, of the target atom located at $\Delta r'$ from the origin O' in subcell A.

Multipole moments of the neighboring subcells further away than subcell C are gathered into the multipole moments of a larger supercell (M2M in Figure 2). The gathered multipole moments are transformed into the local expansion coefficient (M2L) on the larger supercell containing the subcell A. This local expansion coefficient is then decomposed and assigned to subcell A (L2L). Finally, using eq 3, the interactions of the i th atom can be evaluated from the local expansion coefficient on subcell A.

3.2. FMM in the Periodic Boundary Condition. Next, we explain how to calculate the interactions from image cells around the unit cell under the periodic boundary condition. The multipole moment in the image cells is the same as the one in the unit cell M_n^m . Then, the multipole moment from all image cells to the unit cell except for the unit cell itself can be given by

$$\begin{aligned} M_n^m \sum_{\nu \neq 0} R_n^m(r_\nu, \theta_\nu, \varphi_\nu) \\ = M_n^m \left\{ \sum_{\nu \neq 0} \frac{Y_n^m(\theta_\nu, \varphi_\nu) \Gamma(n + 1/2, \kappa^2 r_\nu^2)}{\Gamma(n + 1/2) r_\nu^{n+1}} \right. \\ \left. + \sum_{h \neq 0} \frac{i^n \pi^{n-1/2} Y_n^m(\theta_h, \varphi_h) v_h^{n-2} \exp(-\pi^2 v_h^2 / \kappa^2)}{\Gamma(n + 1/2) V} \right\} \quad (4) \end{aligned}$$

where $(r_\nu, \theta_\nu, \varphi_\nu)$ and $(\nu_h, \nu_h, \varphi_h)$ are the spherical coordinates of the ν th image cell and the reciprocal vector, respectively, and $R_n^m(r_\nu, \theta_\nu, \varphi_\nu)$ is the dimensionless contribution from the image cell ν . κ is a screening parameter similar to the Ewald screening parameter. $\Gamma(n)$ and $\Gamma(n, x)$ are the gamma- and incomplete gamma-function, respectively. In a similar manner to the Ewald method, the interaction is decomposed into two parts, as shown in the right-hand side of eq 4. The first term rapidly decreases to zero with increasing r_ν . The second term is calculated using the Fourier transform. Here, because the number of particles in the cell is just one, that is, the center of the cell, the Fourier transform load is very light. Furthermore, as clearly shown in eq 4, because the angles θ_h and φ_h do not change for the cubic and tetragonal cells throughout the MD calculation, the second term may be calculated and its result communicated to all cells just once at the beginning of the calculation. Even for the fully flexible cell, this small item may be included in the M2M communication. Equation 4 is added to the local expansion coefficient of the unit cell and is used for the interaction calculation of the i th atom in the unit cell. Using the multipole and local expansions stated above, the FMM calculation becomes $O(N)$ compared with $O(N \log N)$ for the FFT-based PME method.

4. DESIGN TO ACHIEVE HIGH PERFORMANCE ON MASSIVELY PARALLEL COMPUTERS

One of the remarkable characteristics of current molecular dynamics simulations is their fine-grained parallelism that can be exploited by parallel computers. This is because of the enormous number of time steps required to achieve the desired results, and thus, the necessity for very short run times per MD step.

The K-computer has 82 944 compute nodes. In this kind of parallel computer, references to global data, which are mostly related to the long-range force calculation, must harm the fine-grained parallelism of the molecular dynamics calculation. Thus, to exploit this computer's capability, efficient new methods are necessary for both data communications and arithmetic operations, together with sophisticated data structures.

In MODYLAS, data structures that enable virtually buffering-free operation in both communications and arithmetic operations are employed. Furthermore, completely localized communication operations by simultaneous localized gather operations between adjacent processes is implemented for communicating both atom coordinates and multipole moments, to enable efficient fine-grained massively parallel processing of molecular dynamics calculations.

4.1. Buffering-Free Data Structure. **4.1.1. Data Structures for Atom Coordinates, Velocities, and Forces.** Figure 4

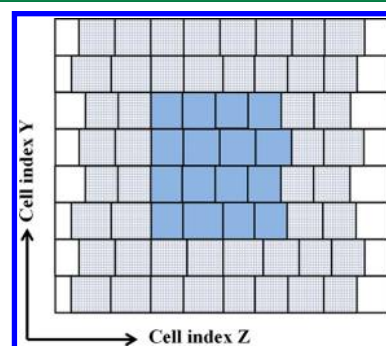


Figure 4. Data array structures for atom coordinates, velocities, and forces. In this example, 4×4 subcells localized in the self-process (blue hatching) are surrounded by subcells transferred from adjacent processes (lattice hatching). Margin areas are placed at both ends in the Z-direction (no hatching).

shows the data structures for atom coordinates, velocities, and forces that we have employed in MODYLAS. The arrays assigned to each MPI process are completely localized to the process, including a vacant area where communicated data are prepared.

Metadata that indicate subcell boundaries and the number of subcell entries have been introduced. The metadata have three-dimensional subcell indices and enable processing of atom coordinates, velocities, and forces for a block of subcells in both arithmetic operations and communications as if they were in three-dimensional fixed size arrays.

Generally, the number of atoms per subcell fluctuates with the motion of atoms. A margin area is reserved as a buffer for the variation in the number of atoms where the Y-directional or X-directional subcell index changes. The cell metadata are always updated according to the motion of atoms from a subcell to a neighbor.

Table 1. Deviation of the Potential Energy and Forces Calculated by the FMM with the Spherical Harmonic Expansion up to the j_{\max} th Order from the Correct Ones by the Ewald Method for an Arbitrarily Chosen Configuration of the PYP Solution^a

j_{\max}	V/kJ mol ⁻¹	σ_x	σ_y	σ_z
4	$-2.302819369602 \times 10^5$	3.0×10^{-4}	3.0×10^{-4}	3.0×10^{-4}
6	$-2.302814155096 \times 10^5$	2.5×10^{-5}	2.5×10^{-5}	2.5×10^{-5}
8	$-2.302814086070 \times 10^5$	2.3×10^{-6}	2.3×10^{-6}	2.3×10^{-6}
10	$-2.302814091104 \times 10^5$	2.5×10^{-7}	2.5×10^{-7}	2.5×10^{-7}

^a $V_{\text{Ewald}} = -2.302814090284 \times 10^5$ kJ mol⁻¹. $\sigma_\alpha = ((\sum_i^N (F_{\alpha,i}^{\text{FMM}} - F_{\alpha,i}^{\text{Ewald}})^2) / (\sum_i^N (F_{\alpha,i}^{\text{Ewald}})^2))^{1/2}$ is the evaluating function for the deviation of force in the α direction ($\alpha = x, y$, and z).

4.1.2. Data Structure for Multipole Moments. To accommodate multilevel multipole moments for domain decomposition on the lowest-level subcell space, in which computation is the most time-consuming, we allow the multipole moment for a single upper-level subcell to be shared by multiple MPI processes when the number of subcells is less than the number of processes in each direction. This implies that redundant operations among processes within the same subcell are required for both arithmetic computation and data communication.

4.2. Fully Localized Adjacent Communication.

4.2.1. Atom Coordinates Communication. The algorithm for data collection from adjacent processes consists of three steps, corresponding to the three directions of communication in the three-dimensional geometry of the model. When each step has been completed, all data from adjacent processes in the direction of communication are collected, and the collected data together with local data are collected in the next step. This strategy minimizes the number of communication steps and the communication data volume. Together with the algorithm, the order of communication is chosen to enable virtually buffering-free communication. The metadata are used to handle atom data as if they were stored in an array of three dimensions corresponding to the model geometry.

4.2.2. Multipole Moment Communication. For the lower levels of the FMM, the same communication strategy as for atom coordinates is employed, when a single supercell is not shared by multiple processes. When a single subcell is shared by multiple processes for upper levels, communications of the same data are made in parallel redundantly among the processes of the single subcell. The communication pair is between processes of the same relative position in two adjacent subcells. When the number of subcells per process is insufficient for M2L operation, which requires four or five adjacent subcells, a bucket-relay method among adjacent subcells is employed.

4.3. Boosting Arithmetic Performance by Localizing Hot Spot Data. A key to achieving high arithmetic performance at a hot spot is to utilize the cache efficiently. It is essential to minimize the traffic between main memory and the level-2 cache and that between the level-2 cache and level-1 cache. In MODYLAS, special loop orderings are implemented to achieve an extremely low cache miss rate at two hot spots, calculations of interatomic pairwise additive interactions and M2L operation of multipole moments. The idea for higher cache efficiency is to perform almost all operations that refer to the data once they are loaded into the cache. For this purpose, in the pairwise additive interaction, atom data are treated by a block of multiple subcells. Loop ordering is such that a block in the cache can be reused for all the operations referring to the block. The use of atom pair lists is avoided by using the metadata. For the M2L operation for multipole moments, the

target of reuse of data in the cache is the M2L operation matrix. All operations that refer to the matrix of the same relative subcell address are carried out before switching to the next relative subcell address.

5. BENCHMARK TEST

In this section, we give a brief description of the benchmark MD calculation on the K-computer. A very large molecular system was adopted for the benchmark test because we are interested in such very large systems as viruses, liposomes, and assemblies of micelles and proteins. The hardware system, the K-computer, is also very large.

5.1. Benchmark MD Calculation. For the benchmark test in the present study, we used photoactive yellow proteins (PYP) in water. This system includes all the factors required by most of the MD calculations used in the scientific and technological investigations. The system consists of 512 PYP molecules, 3 005 952 water molecules, and 2560 counterions, for a total of 9 999 872 atoms. All atoms in the system interact with each other, taking account of the long-range Coulombic force, although the system is just for the benchmark test. We call this system our 10 million-atom system. CHARM22 with CMAP and TIP3P were adopted for the force fields of PYP and water, respectively. The cutoff distance for the short-range LJ potential was 1.2 nm, and the spherical harmonic expansion up to the fourth order was taken into account to evaluate Coulombic interactions. The lengths of chemical bonds relevant to the hydrogen atoms were constrained by SHAKE. The test was performed using the NVE ensemble. When the calculation was done to form an NPT ensemble, the calculation time increased by 10–20%, mostly because of ROLL and relevant global communications of the kinetic energy and virial. The MD unit cell was decomposed into 64^3 subcells, that is, six levels of the FMM. Because the FMM deals with the interactions with distant molecules, a multiple time-step method was applied in which direct calculations of LJ interaction and Coulombic interaction with the molecules in the near-neighbor subcells were done every time step, $\Delta t = 2$ fs, while FMM calculations for the Coulombic interaction with the molecules in distant subcells were done every four time steps, $4\Delta t$, and the FMM calculations with respect to the molecules in the far distant subcells were done every eight time steps, $8\Delta t$.

5.2. Benchmark System. The K-computer was used in this benchmark test. Its directly connected network enables simultaneous communication with up to four directions at a time. Furthermore, the compute node is equipped with single-instruction multiple-data (SIMD) operation of a floating point multiply-and-add unit, and a reciprocal approximation instruction for $1/x$ and $1/\sqrt{x}$. The theoretical peak performance is 10.61 Pflops. In the present performance test, 65 536 ($\approx 2^{16}$) nodes (about 80% of the total system) were used.

6. RESULTS AND DISCUSSION

6.1. Long-Range Force by FMM Calculation. Long-range forces on an atom evaluated by the present FMM are used directly to solve the equation of motion for the atom. Thus, accuracy in the calculation of the long-range forces is essential for high-quality MD calculations. In this section, the accuracy of the FMM forces calculated for our 10 million-atom system is examined, comparing them with values calculated very accurately with a different method. To obtain the reference values, we adopted the Ewald method with a sufficiently large convergence parameter in real space ($0.7 \times 10^{10} \text{ m}^{-1}$ for $r_{\text{cut}} = 1.2 \text{ nm}$; i.e., fast convergence of the error function) and an excessively fine reciprocal space ($h_{\text{max}}^2 = 350\,000$) at very high calculation cost. Twelve figures of the calculated Ewald forces do not vary as a function of h_{max}^2 for this large h_{max}^2 ; we chose these values as our reference.

Table 1 compares the Coulombic total potential energy and forces for an arbitrarily chosen configuration of the PYP solution calculated by the FMM with the very accurate values from the Ewald method. The evaluating function σ_α for the deviation of a force from the correct value may be defined by

$$\sigma_\alpha = \sqrt{\frac{\sum_i^N (F_{\alpha,i}^{j_{\text{max}}} - F_{\alpha,i}^{\text{Ewald}})^2}{\sum_i^N (F_{\alpha,i}^{\text{Ewald}})^2}} \quad (5)$$

where $F_{\alpha,i}^{j_{\text{max}}}$ and $F_{\alpha,i}^{\text{Ewald}}$ are the forces acting on atom i in the direction $\alpha = x, y$, and z calculated by the FMM with the expansion up to the j_{max} th order and the Ewald method, respectively. N is the number of atoms in the system (i.e., 9 999 872 in the present calculation). Figure 5 also shows the

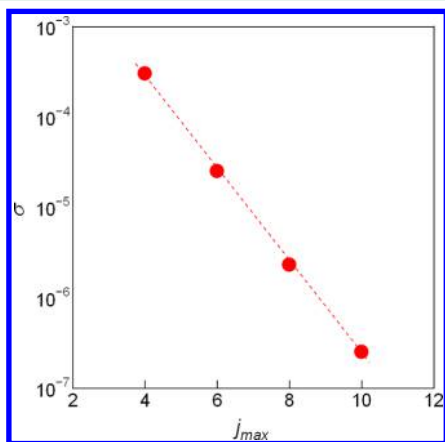


Figure 5. Deviation σ ($= \sigma_x = \sigma_y = \sigma_z$) of the forces calculated by the FMM from the correct ones by the Ewald method for an arbitrarily chosen configuration of the 10 million-atom system, the PYP solution.

convergence of the calculated forces to the correct ones represented by σ_α as a function of the order of the spherical harmonic expansion. The table and figure start with the fourth expansion because the Ewald method used for the calculation of interaction with the multipoles of the total MD unit cell requires at least the fourth-order calculation. The table clearly shows that the FMM total potential energy agrees well with the correct one by the Ewald method to six figures for the expansion up to the fourth order, and the error decreases logarithmically such that the value agrees with the correct value to nine figures for the expansion up to the 10th order. Deviations of the forces acting on the atoms in three directions

are almost the same and are in the order of 10^{-4} for the fourth-order expansion. The deviation decreases logarithmically such that it is as small as 10^{-7} for the 10th-order expansion. These results clearly show that the FMM energy and forces converge well to the correct values as the order of expansion increases, and the fourth-order expansion is satisfactory for ordinary MD calculations.

In Figure 6, the total Hamiltonian H of the PYP system is shown as a function of MD time, where the long-range forces

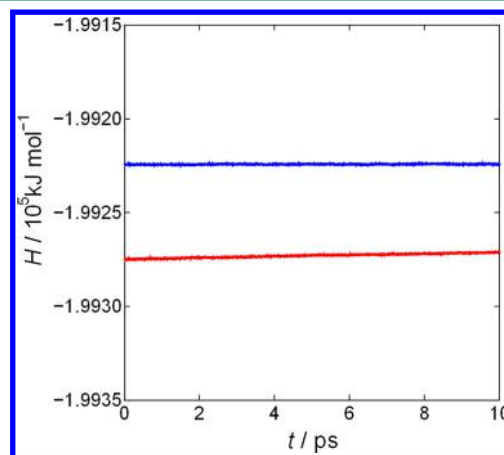


Figure 6. Total Hamiltonian H for the 10 million-atom system, the PYP solution, where the forces were calculated by the FMM with spherical harmonic expansion up to the fourth (lower line) order and sixth order (upper line).

were evaluated by the FMM with the expansion up to the fourth order and sixth order. In these MD calculations, the motions caused by the intramolecular forces, such as stretching and bending, were solved with the short time step, $\Delta t/4$. As clearly shown in the figure, the conservation of the Hamiltonian is good. This implies that the long-range forces are evaluated by the FMM with satisfactory precision for ordinary MD calculations.

6.2. Overall Performance. Figure 7 shows the strong scalability of the measured calculation time per MD step Δt , using from 64 nodes (512 cores) to 65 536 nodes (524 288 cores). The figure also shows the acceleration with respect to the 64-node calculation. The scalability of the degree of parallelism is excellent, better than 90%, with fewer than about 10 000 compute nodes. For $N_{\text{node}} = 65\,536$, the plot somewhat deviates from the ideal acceleration. However, the calculation time is still as short as about 5 ms. This demonstrates that the present performance belongs to one of the highest classes of the MD calculation for very large molecular systems. The acceleration rate for the 65 536-node calculation is 412 with respect to the 64-node calculation, the parallelization efficiency being still about 40% in spite of the significant increase in parallelized nodes, $65\,536/64 = 1024$. The performance is in good contrast to that by the PME method, where the efficiency becomes very low for degree of parallelism higher than 256 and 512.^{21,22}

The high performance is due to the extremely good main arithmetic operations as well as the excellent parallelization reported below.

6.3. Analysis of the Performance. The performance rate measured for the direct interaction calculation using the 65 536 nodes was 32.0% of the theoretical peak performance for the 10

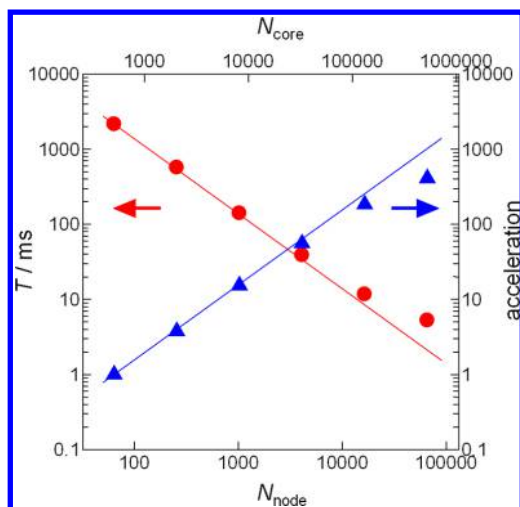


Figure 7. Measured overall calculation time per MD step (Δt) for the 10 million-atom system, the PYP solution, and the acceleration ratio with respect to the 64-node calculation as a function of the degree of parallelism, N_{node} .

million-atom system. In contrast, the performance for the whole FMM calculation was a little lower, 14.3%. However, the small amount of the FMM calculation compared with the direct interaction calculation does not reduce the overall performance so much. According to a performance analysis run for small MD system, the SIMD instruction rates were 79% for direct force calculation and 83% for the M2L operation of the lowest-level moments in the FMM. Furthermore, following the special loop ordering explained in section 4.3, level-1 data cache miss rates per load/store instruction are extremely small, 1.22%, for the direct force calculation, and 1.38% for lowest-level M2L calculation of multipole moments. The level-2 cache miss rates per load/store instruction are 0.001% and 0.12% for the former and the latter, respectively. These cache miss rates will be closely maintained, at least, even for the 80 million-atom system. The excellent arithmetic operations are responsible for the very high performance of the overall MD calculations stated.

In Figure 8, the communication time per one MD step, Δt , is presented for coordinates, multipole moments, and atoms crossing the boundaries of subcells. It scales well for thousands of compute nodes, though not for tens of thousands of nodes.

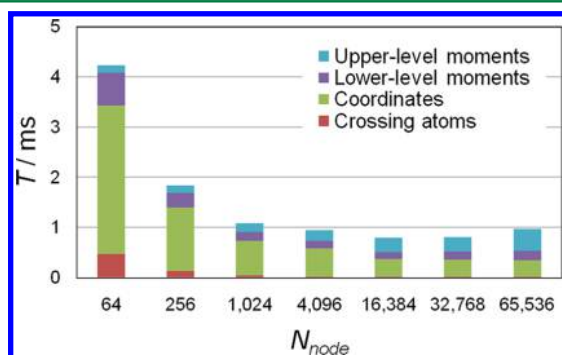


Figure 8. Measured communication time of the upper-level moments (blue), the lower-level moments (purple), the coordinates (green), and the atoms crossing the boundary of subcells (brown) per MD step (Δt) for the 10 million atom system, the PYP solution, as a function of degree of parallelism, N_{node} .

However, it is notable that the time required for the communication is very short (<1 ms) and does not increase much, even when the number of compute nodes becomes very large ($>10\,000$). This excellent performance of the communications operations is obtained from both the virtually buffering-free data structure and the fully localized adjacent communication operations described in section 4. The present communication time, about 1 ms, is negligibly short compared with the arithmetic operations when the number of nodes is smaller than 10 000, as shown later.

In the figure, the coordinate communication as well as the lower-level multipole moment communication seems to be almost constant when the degree of parallelism becomes very high. This is caused by a combination of the inevitably increasing number of communications and the decreasing data volume of each communication. The former increases the communication time because the aggregate MPI communication setup latency is proportional to this, while the latter contributes to the decrease in actual communication transmissions. The upper-level multipole moment communication increases slightly for tens of thousands of nodes. This is because it includes communications covering a wide area (i.e., large supercells), and hence, it suffers from communication congestion on a network route. However, in this case, too, the time required for the communication is in the order of 10^2 μs .

It is interesting to consider again the slightly sluggish performance found when the number of compute nodes exceeded 10 000. Figure 9 presents the calculation time for the

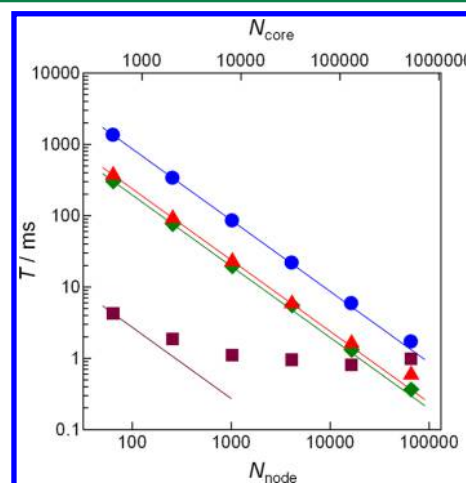


Figure 9. Measured direct calculation time of the pairwise additive forces (blue), the FMM calculation (orange), the intramolecular forces (green), and the communication time (brown) per MD step (Δt) for the 10 million-atom system, the PYP solution, as a function of degree of parallelism, N_{node} .

pair interactions with the molecules located in the first and second nearest subcells (circles), the FMM calculation time (triangles), the calculation time for the intramolecular forces (diamonds), and the communication time (squares) per MD time step, Δt . As the figure clearly shows, the former three scale very well for $N_{\text{node}} < 10\,000$: the plots show almost ideal behavior on the straight line. However, they deviate a little from the straight line when the number of compute nodes becomes very large, $N_{\text{node}} > 10\,000$. This is one reason why the total

performance does not show the ideal parallelization efficiency for very large numbers of nodes.

Furthermore, the communication time also becomes important in the total calculation time. It may be negligible for low degrees of parallelism where the arithmetic operation times are relatively very long. However, in contrast to the almost constant communication time, the total calculation time becomes very short as the number of nodes becomes large, and the communication time occupies a nonnegligible portion of the total time. This is another reason for the degradation of the parallelization efficiency with very large numbers of nodes.

Other factors such as load imbalance and incidental system noise can be additional sources of degradation of efficiency with very short calculation times using very large numbers of nodes, because the values presented here are averaged over the nodes. For $N_{\text{node}} = 65\,536$, the former increases the calculation time by 0.8 ms and the latter also by 0.8 ms, resulting in the total calculation time of 5 ms.

However, we would like to stress that the efficiency is still satisfactory even for $N_{\text{node}} = 65\,536$ to give excellent performance, 5 ms per MD step, for a 10 million-atom system. For larger molecular systems (e.g., 80 million-atom systems), degradation was not found even for the very high degree of parallelism, $N_{\text{node}} = 65\,536$.

7. CONCLUSIONS

We have developed a highly scalable general-purpose MD simulation program, MODYLAS, in which long-range forces are evaluated rigorously using the FMM. The method is free from FFT calculations, which have severely hindered massively parallel supercomputers. Several new methods have also been developed for extremely fine-grained parallelism of the MD calculation. The virtually buffering-free communications and arithmetic operations, the minimal communication latency algorithm, and the parallel bucket-relay algorithm for the upper-level FMM communications realize excellent scalability. New arithmetic algorithms have also been developed. In particular, the methods for blockwise arithmetic operation for the pair-interaction calculations are successful in avoiding data reload, attaining very small level-1 and level-2 cache miss rates of less than 2% and 0.1%, respectively. Together with the excellent communications up to very high degrees of parallelism, MODYLAS enables investigations of 10 million-atom real systems such as viruses, liposome, assemblies of proteins and micelles, and polymers, for which a 100 ns-long MD calculation may be completed in 3 days (5 ms/step). We believe that MODYLAS will open a new frontier in computational physical, chemical, and biological sciences at a molecular level. For example, long-time all-atom MD calculations are in progress for viral systems.

AUTHOR INFORMATION

Corresponding Author

*E-mail: okazaki@apchem.nagoya-u.ac.jp.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Next Generation Super Computing Project, Nanoscience Program, and by TCCI/CMSI in the Strategic Programs for Innovative Research, MEXT, Japan. Main results were obtained by early access to the

K computer at the RIKEN Advanced Institute for Computational Science, and the calculations were partly performed at the Research Center for Computational Science, Okazaki, Japan, the Center for Computational Sciences, University of Tsukuba, Tsukuba, Japan, and the Supercomputer Center, the Institute for Solid State Physics, the University of Tokyo. The authors also thank to Prof. K. Yasuoka at Keio University, Prof. T. Narumi at the University of Electro-Communications, Dr. A. Kawai at the K&F Computing Research Co., Dr. W. Shinoda at the AIST, and Dr. H. Watanabe at University of Tokyo for their valuable advice. The MODYLAS will be opened to any researcher. For more information, contact the corresponding author.

REFERENCES

- (1) Ajima, Y.; Sumimoto, S.; Shimizu, T. Tofu: A 6D Mesh/Torus Interconnect for exascale computers. *Computer* **2009**, *42*, 36–40.
- (2) Toyoshima, T. ICC: An interconnect controller for the Tofu interconnect architecture. *Hot Chips: A Symposium on High Performance Chips*, Stanford University, Aug. 22–24, 2010.
- (3) Fukushima, T.; Taiji, M.; Makino, J.; Ebisuzaki, T.; Sugimoto, D. A highly parallelized special-purpose computer for many-body simulations with an arbitrary central force: MD-GRAPE. *Astrophys. J.* **1996**, *468*, 51–61.
- (4) Komeiji, Y.; Yokoyama, H.; Uebayashi, M.; Taiji, M.; Fukushima, T.; Sugimoto, D.; Takata, R.; Shimizu, A.; Itsukashi, K. A high performance system for molecular dynamics simulation of biomolecules using a special-purpose computer. In *Proceedings of the Pacific Symposium on Biocomputing '96*; Hunter, L., Klein, T. E., Eds.; World Scientific Publishing: Singapore, 1995; p 472.
- (5) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Kolossváry, I.; Klepeis, J. L.; Layman, T.; McLeavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C. Anton: A special-purpose machine for molecular dynamics simulation. *Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA '07)*; ACM: New York, 2007.
- (6) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large system. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (7) Greengard, L. F. *The Rapid Evaluation of Potential Fields in Particle Systems*; MIT Press: Cambridge, MA, 1988.
- (8) Figueirido, F.; Levy, R. M.; Zhou, R.; Berne, B. J. Large scale simulation of macromolecules in solution: Combining the periodic fast multiple method with multiple time step integrators. *J. Chem. Phys.* **1997**, *106*, 9835–9849.
- (9) Yonezawa, A.; Watanabe, T.; Yokokawa, M.; Sato, M.; Hirao, K. Advanced Institute for Computational Science (AICS) Japanese National High-Performance Computing Research Institute and its 10-petaflops supercomputer K. *International Conference for High Performance Computing, Networking, Storage, and Analysis*; IEEE: Washington, DC, 2011.
- (10) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (11) Hoover, W. G. Constant pressure equations of motion. *Phys. Rev. A* **1986**, *34*, 2499–2500.
- (12) Andersen, H. C. Molecular dynamics simulation at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
- (13) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equation of motion of a system with constraints: Molecular dynamics of n -alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (14) Andersen, H. C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculation. *J. Comput. Phys.* **1983**, *52*, 24–34.

- (15) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithm. *J. Chem. Phys.* **1994**, *101*, 4117–4189.
- (16) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117–1157.
- (17) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
- (18) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ , ψ , and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (19) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins*. **2006**, *65*, 712–725.
- (20) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem.* **2001**, *105*, 6474–6487.
- (21) Gruber, C. C.; Pleiss, J. Systematic benchmarking of large molecular dynamics simulations employing GROMACS on massive multiprocessing facilities. *J. Comput. Chem.* **2011**, *32*, 600–606.
- (22) Mei, C.; Sun, Y.; Zheng, G.; Bohm, E. J.; Kale, L. V.; Philips, J. C.; Harrison, C. Enabling and scaling biomolecular simulations of 100 million atoms on petascale machines with a multicore-optimized message-driven runtime. *International Conference for High Performance Computing, Networking, Storage, and Analysis*; IEEE: Washington, DC, 2011.