John von Neumann Institute for Computing

A High-Speed Special-Purpose Computer for Molecular Dynamics Simulations: MDGRAPE-3

Tetsu Narumi, Yousuke Ohno, Noriaki Okimoto, Atsushi Suenaga, Ryoko Yanai, Makoto Taiji

published in

NIC Workshop 2006,
From Computational Biophysics to Systems Biology,
Jan Meinke, Olav Zimmermann,
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,
ISBN-13: 978-3-9810843-0-6, pp. 29-36 , 2006.

http://www.fz-juelich.de/nic-series/volume34

A High-Speed Special-Purpose Computer for Molecular Dynamics Simulations: MDGRAPE-3

Tetsu Narumi, Yousuke Ohno, Noriyuki Futatsugi, Noriaki Okimoto, Atsushi Suenaga, Ryoko Yanai, and Makoto Taiji

High Performance Molecular Simulation Team,
Computational and Experimental Systems Biology Group,
Genomic Sciences Center, RIKEN
61-1-1F, Ono-cho, Tsurumi, Yokohama, Kanagawa, 230-0046, JAPAN
E-mail: narumi@gsc.riken.jp

We developed a high-speed computer, MDGRAPE-3, which is dedicated to perform molecular dynamics simulations. It is a special-purpose computer for force calculation between particles and has become the first Pflops machine. Coulomb and van der Waals force calculations are accelerated by the MDGRAPE-3, and the other calculations are performed on a usual general-purpose computer. By using this architecture, we can get higher performance and 10 times lower cost, size, and power consumption per performance than those of general-purpose computers. MDGRAPE enables us to perform precise molecular dynamics simulations, because explicit water molecules and non-cutoff Coulomb interaction can be efficiently accelerated by this hardware.

1 Introduction

One of the main target of the recent supercomputers is molecular dynamics (MD) simulations, which require huge computational power. To satisfy this demand, cluster of computers is often used because its cost per performance is very low. Even the high-end supercomputers are made of commodity based PCs. For example, 20-25% of the top 50 supercomputers are PC clusters[1]. However, PC clusters have many disadvantages; too big and power consuming. One of the reasons why only one PC cluster is in the top 10 supercomputers is because of its size and power consumption level.

One approach is to develop a computer which is specially packed in a small form factor. IBM has developed the Blue Gene/L[2], which is the fastest in the world for the general-purpose computer with relatively small and low power consumption.

We took another approach to reach the highest performance: Special-purpose computer for particle simulations. GRAPE[3–5] is a series of special-purpose computers for astrophysical N-body and MD simulations, etc. MDGRAPE-3 is the latest computer for MD simulations[6]. MDGRAPE-3 becomes the first Pflops machine, and its cost, size and power consumption per performance is 10 times smaller than those of any other computers.

In this paper, hardware, software, and performance of MDGRAPE-3 system are presented as well as the short report of MD simulations with our hardware. Hardware and software of MDGRAPE-3 system are shown in Secs. 2 and 3 respectively. Performance is presented in Sec. 4. MD simulation results and conclusion are described in Secs. 5 and 6 respectively.

2 Hardware of MDGRAPE-3 System

The GRAPE is a series of application specific processor designs, which is specially built to accelerate the inter-particle force calculation. The GRAPE was originally developed for accelerating the gravitational interactions, while MD simulations are also supported in the case of GRAPE-2A[7], MD-GRAPE[8,9] and MDM (MDGRAPE-2)[10–12]. The main difference in hardware which supports MD and which does not is that the arbitrary central force can be calculated with MD-supported machines instead of the gravitational force.

Machines other than GRAPE have been developed with similar idea: such as Delft Molecular Dynamics Processor[13], ATOMS[14], FASTRUN[15] and so on. However, these machines were not so successful compared with GRAPE because their hardwares were complicated. MDGRAPE is simple because it only accelerates the long range interactions (Coulomb and van der Waals). Other calculations are performed in the host computer which is connected to MDGRAPE. In MD simulations, long range interactions dominate the total calculation. Therefore, the simple hardware can accelerate MD simulations more effectively.

To reach high performance, we developed a processor, MDGRAPE-3 chip, which is dedicated to MD simulations. It has 180 or 216 Gflops equivalent performance with the clock frequency of 250MHz or 300MHz when the effective number of floating-point operations of a pairwise force calculation is assumed to be 36. In this assumption, we counted division and square root operations as ten floating-point operations respectively. There are two main reasons why the MDGRAPE-3 chip has high performance. One reason is that MDGRAPE-3 chip has 20 special pipelines, each of which calculates pairwise interaction between atoms. Many hundreds of arithmetic units and their data path in the pipelines are hard wired. Therefore, all the arithmetic units work every clock cycle basically, and it makes highly efficient calculation possible. On the other hand, conventional CPU, such as Pentium, can perform only several operations per clock cycle. Another reason is that all the pipelines can share one memory due to the parallelism of particle interaction. This lower bandwidth of memory enables simpler hardware, which means it is easy to parallelize to get higher performance.

Figure 1(left) shows the MDGRAPE-3 board which has 12 MDGRAPE-3 chips. The memory which stores particle positions etc. is integrated in the MDGRAPE-3 chip, and

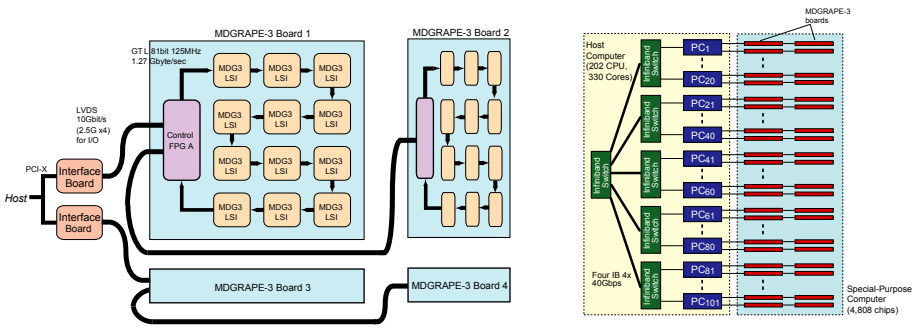

Figure 1. Block diagram of the MDGRAPE-3 board, which has 12 MDGRAPE-3 chips connected in a daisy-chain and an FPGA (left). Block diagram of the MDGRAPE-3 system, which is consisted of the PC cluster with 101 nodes and 402 MDGRAPE-3 boards (right).

System	Peak speed (Gflops)	Cost/speed ($/Gflops)	Power/speed (Watt/Gflops)	Size/speed (liter/Gflops)
MDGRAPE-3	1,000,000	**9**	**0.2**	**0.03**
Blue Gene/L	360,000	280	4	0.3
PC (Pentium D 3.0GHz)	12	80	25	3

Table 1. Advantages of MDGRAPE-3 against general-purpose computers. The MDGRAPE-3 is superior to general-purpose computers in all of the three points which are cost/speed, power/speed, and size/speed as underlined.

the board is also simple. It is connected to the host computer via interface board. Figure 1(right) shows the total system. It is composed of a PC cluster with 101 nodes, each of them has four MDGRAPE-3 boards. Total number of CPU cores is 330, because we used 64 nodes of dual Xeon 5150 and 37 nodes of dual Xeon 3.2D GHz CPUs.

MDGRAPE-3 has many advantages compared to a conventional computer. Table 1 compares special- and general-purpose computers in four points: Peak speed, cost per speed, power consumption per speed, and size per speed. We chose Blue Gene/L and a PC for general-purpose computers, because Blue Gene/L is the fastest supercomputer, one of the lowest power, and smallest size computer per performance, and PC is one of the most cost effective computer. MDGRAPE-3 is superior to general-purpose computers in all four categories above. It has 10 times lower cost, size, and power consumption per performance than those of general-purpose computers.

3 Software of MDGRAPE-3 System

To get the benefit of MDGRAPE-3, the application program must be modified to fit our hardware. We have already ported AMBER[16] and CHARMM[17] programs, though some limitations exist. Users just need to add keyword to use MDGRAPE-3 in input files.

Other application can be accelerated with our APIs (Application Programming Interfaces). Table 2 is one of the APIs to calculate the Coulomb force between atoms. Software for the MDGRAPE-3 can be downloaded from the web site[18]. We support C and Fortran APIs and Linux operating system.

```
void MR3calccoulomb(double x[], int n, double q[],double rscale, int tblno,
                    double size, int pflag, int flag, double force[])
 x ------ Positions of atoms.   n ----- Number of atoms.
 q ------ Charges of atoms.     tblno - Potential type between atoms.
 xmax --- Simulation box size.  flag -- Special flag to control calculation.
 rscale - Alpha parameter for real-space of Coulomb of the Ewald method.
 pflag -- Flag to specify periodic boundary condition or not.
 force -- Calculated forces (output).
```

Table 2. Example of APIs in C to calculate the Coulomb force between atoms

Number of atoms	3,339	4,953	13,101
Without MDGRAPE-3 card (sec)	346.73	844.52	9231.70
With MDGRAPE-3 card (sec)	6.38	7.33	34.31
Acceleration	**54**	**115**	**269**

Table 3. Performance of the MDGRAPE-3 card

4 Performance of MDGRAPE-3 System

Table 3 summaries the performance of MDGRAPE-3. We used 2-chip version of MDGRAPE-3 card, which is commercially available from the venture company[19]. We compared the calculation time of the 1,000 steps of 'sander' program in AMBER8 with and without MDGRAPE-3 card. The calculated system has a Scytalone dehydratase (2,715 atoms) and water molecules. Coulomb and van der Waals interactions are calculated without cutoff under free boundary condition. Pentium D 3.0 GHz (only single core is used) and Pentium 3.2 GHz CPUs are used with and without MDGRAPE-3 calculation respectively. We got more than 100 times acceleration.

For PME (Particle Mesh Ewald) method simulation, the acceleration is not so high because the wavenumber-space part of the Coulomb calculation becomes the bottleneck. FFT (Fast Fourier Transform), which is used for wavenumber-space, cannot be accelerated with MDGRAPE-3. With newer method such as IPS[20] (Isotropic Periodic Sum), we will be able to accelerate more efficiently under periodic boundary condition in the near future.

5 MD Simulation Results

In this section, we briefly present MD simulation results with previous MDGRAPE-2 hardware. First example is RNA polymerase II[22] which is the fundamental and important enzyme in the transcription process. As pre-translocation state has not yet been determined by the X-ray crystallographic study, we first constructed a model structure of the Pol II elongation complex with the 50 base pairs of DNA-24 bases of RNA, including the unwound bubble of DNA. Then we got its reliable structure by 2 ns of MD simulation [Fig.

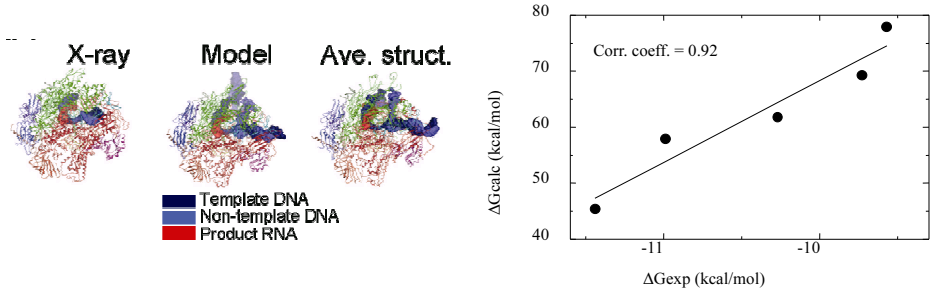

Figure 2. X-ray[21], initial model, and average structures of the Pol II elongation complex are shown from left to right (left). Correlation between calculated and experimental binding free energy between SH2 domain of growth factor receptor binding protein 2 and ErbB receptor-derived phosphotyrosyl peptides (right).

2(left)]. It was a big system with 450 k atoms in total, and MDGRAPE hardware enabled us to perform accurate simulation with explicit water molecules and non-cutoff of Coulomb and van der Waals interaction. We also performed a short simulation with a targeted MD simulation approach[23] and found that the conformational change of a loop in the Pol II, fork loop 1, couples with the unidirectional movement of the Pol II along DNA.

Second example is free energy calculation between the SH2 domain of growth factor receptor binding protein 2, Grb2, and ErbB receptor-derived phosphotyrosyl peptides[24]. Binding free energies for nine phosphotyrosyl peptides were calculated by using the MM-PBSA (Molecular Mechanics Poisson-Boltzmann Surface Area) continuum solvent method[25], and we got a good correlation coefficient of 0.92 between simulated and experimental data of surface plasmon resonance analysis [Fig. 2(right)]. Such precise binding free energy calculation would be useful for *in silico* drug screening because conventional docking programs tend to be inaccurate because they often treat proteins as rigid and water molecules as implicit models. With full MDGRAPE-3 system, over 2,000 of simulations, each of which is 2 ns long and has 13 k atoms, can be performed per day.

6 Concluding Remarks

We have described the MDGRAPE-3, special-purpose computer for MD simulations. It is the first Pflops computer, and has 10 times lower cost, size, and power consumption per performance than those of general-purpose computers. Our architecture of special-purpose computer would be promising approach in the near future because power consumption and size are becoming more and more important in conventional processors. We can perform precise MD simulations with our hardware because it efficiently accelerates the simulation with explicit water molecules and non-cutoff interactions.

Acknowledgments

This work was partially supported by Research Revolution 2002 “Protein 3000 Project”, operated by the Ministry of Education, Culture, Sports, Science and Technology, Japan. Authors thank Drs Yoshiyuki Sakaki, Tomoya Ogawa, Shigeyuki Yokoyama, and Toshikazu Ebisuzaki for their support to this project. Authors also thank Japan SGI Inc. for the system support. In particular, we would like to thank Masahiro Yamaguchi, Hiroshige Tatsu, and Junya Shinohara for their help with the system installation. Authors thank Intel Corporation for providing their early processors and system support for them.

References

1. `http://www.top500.org/list/2005/11/100`
2. `http://www.research.ibm.com/bluegene/index.html`
3. D. Sugimoto, Y. Chikada, J. Makino, T. Ito, T. Ebisuzaki, and M. Umemura, *A Special Purpose Computer for Gravitational Many-body Problems*, Nature **345**, 33-35 (1990).
4. J. Makino and M. Taiji, *Scientific Simulations with Special-Purpose Computers: The Grape Systems*, John Wiley & Sons, (1998).

5. P. Hut, and J. Makino, *Computational Physics - Astrophysics on the GRAPE Family of Special-Purpose Computers*, Science **283**, 5401 (501-505)1999.
6. M. Taiji, T. Narumi, Y. Ohno, N. Futatsugi, A. Suenaga, N. Takada, and A. Konagaya, *Protein Explorer: A Petaflops Special-Purpose Computer System for Molecular Dynamics Simulations*, in Proceedings of SC2003, CDROM, Phoenix, (2003).
7. T. Ito, J. Makino, T. Fukushige, T. Ebisuzaki, S. K. Okumura, and D. Sugimoto, *A Special-Purpose Computer for N-Body Simulations: GRAPE-2A*, Publications of the Astronomical Society of Japan **45**, 339-348 (1993).
8. M. Taiji, T. Fukushige, J. Makino, T. Ebisuzaki, and D. Sugimoto, *MD-GRAPE: A Parallel Special-Purpose Computer System for Classical Molecular Dynamics Simulations*, in Proceedings of the 6th Joint EPS-APS international conference on Physics Computing, Geneva, 200-203 (1994).
9. T. Fukushige, M. Taiji, J. Makino, T. Ebisuzaki, and D. Sugimoto, *A Highly-Parallelized Special-Purpose Computer for Many-body Simulations with an Arbitrary Central Force: MD-GRAPE*, Astrophys. J. **468**, 51-61 (1996).
10. T. Narumi, R. Susukita, T. Ebisuzaki, G. McNiven, and B. Elmegreen, *Molecular Dynamics Machine: Special-purpose Computer for Molecular Dynamics Simulations*, Mol. Sim. **21**, 401-415 (1999).
11. T. Narumi, A. Kawai, and T. Koishi, *An 8.61 Tflop/s Molecular Dynamics Simulation for NaCl with a Special-Purpose Computer: MDM*, in Proceedings of SC2001, CDROM, Denver, (2001).
12. R. Susukita, T. Ebisuzaki, B. G. Elmegreen, H. Furusawa, K. Kato, A. Kawai, Y. Kobayashi, T. Koishi, G. D. McNiven, T. Narumi, and K Yasuoka, *Hardware Accelerator for Molecular Dynamics: MDGRAPE-2*, Computer Physics Communications **155**, 115-131 (2003).
13. A. F. Bakker and C. Bruin, in *Special Purpose Computers*, edited by B. J. Alder (Academic Press, San Diego), 183-232, (1988).
14. A. F. Bakker, G. H. Gilmer, M. H. Grabow, and K. Thompson, *A Special Purpose Computer for Molecular Dynamics Calculations*, J. Comput. Phys. **90**, 313-335 (1990).
15. R. Fine, G. Dimmler, and C. Levinthal, *FASTRUN: A Special Purpose, Hardwired Computer for Molecular Simulation*, Proteins **11**, 242-253 (1991).
16. D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman., M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, P. A. Kollman, Amber 8.0, University of California San Francisco 2004.
17. B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus, *CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations*, Comp. Chem. **4**, 87-217 (1983).
18. `http://mdgrape.gsc.riken.jp`
19. `http://www.peta.co.jp`
20. Xiongwu Wu and Bernard R. Brooks, *Isotropic periodic sum: A method for the calculation of long-range interactions*, J. Chem. Phys. **122**, 044107 (2005).
21. H. Kettenberger, K.-J. Armache, and P. Cramer, *Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS*, Mol. Cell **22**, 955965 (2004).

22. A. Suenaga, N. Okimoto, N. Futatsugi, Y. Hirano, T. Narumi, Y. Ohno, R. Yanai, T. Hirokawa, T. Ebisuzaki, A. Konagaya, and M. Taiji, *Structure and dynamics of RNA polymerase II elongation complex*, Biochemical and Biophysical Research Communications **343**, 90-98 (2006).
23. C. Guilbert, D. Perahia, L.A. Mouawad, *A method to explore transition paths in macromoleculesapplications to hemoglobin and phosphoglycerate kinase*, Comput. Phys. Comm. **91**, 263273 (1995).
24. A. Suenaga, , M. Hatakeyama, , M. Ichikawa, X. M. Yu, N. Futatsugi, T. Narumi, K. Fukui, T. Terada, M. Taiji, M. Shirouzu, S. Yokoyama, and A. Konagaya, *Molecular Dynamics, Free Energy, and SPR Analyses of the Interactions Between the SH2 Domain of Grb2 and ErbB Phosphotyrosyl Peptides*, Biochemistry **42**, 5195-5200 (2003).
25. P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Hou, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham III,*Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models*, Acc. Chem. Res. **33**, 889-897 (2000).