

GRAPE Accelerators

Junichiro Makino^{1,2,3}

¹Division of Theoretical Astronomy, National Astronomical Observatory of Japan, 2-21-1
Osawa, Mitaka-shi, Tokyo 181-8588.

²Center for Computational Astrophysics, National Astronomical Observatory of Japan, 2-21-1
Osawa, Mitaka-shi, Tokyo 181-8588

³School of Physical Sciences, Graduate University of Advanced Study (SOKENDAI), 2-21-1
Osawa, Mitaka-shi, Tokyo 181-8588
email: makino@cfca.jp

Abstract. I'll overview the past, present, and future of the GRAPE project, which started as the effort to design and develop specialized hardware for gravitational N -body problem. The current hardware, GRAPE-DR, has an architecture quite different from previous GRAPEs, in the sense that it is a collection of small, but programmable processors, while previous GRAPEs had hardwired pipelines. I'll discuss pros and cons of these two approaches, comparisons with other accelerators and future directions.

Keywords. methods: n -body simulations — methods: numerical

1. Introduction

In many simulations in astrophysics, it is necessary to solve gravitational N -body problems. In some cases, such as the study of formation of galaxies or stars, it is important to treat non-gravitational effects such as the hydrodynamical interaction, radiation, and magnetic fields, but in these simulations calculation of gravity is usually the most time-consuming part.

To solve the gravitational N -body problem, one needs to calculate the gravitational force on each body (particle) in the system from all other particles in the system. There are many ways to do so, and if relatively low accuracy is sufficient, one can use the Barnes-Hut tree algorithm (Barnes & Hut 1986) or FMM (Greengard and Rokhlin 1987). Even with these schemes, the calculation of the gravitational interaction between particles (or particles and multipole expansions of groups of particles) is the most time-consuming part of the calculation. Thus, one can greatly improve the speed of the entire simulation, just by accelerating the speed of the calculation of particle-particle interaction. This is the basic idea behind GRAPE computers.

The basic idea is shown in figure 1. The system consists of a host computer and special-purpose hardware, and the special-purpose hardware handles the calculation of gravitational interaction between particles. The host computer performs other calculations such as the time integration of particles, I/O, and diagnostics.

2. History

GRAPE Project was started in 1988. The first machine completed, the GRAPE-1 (Ito *et al.* 1990), was a single-board unit on which around 100 IC and LSI chips were mounted and wire-wrapped. The pipeline processor of GRAPE-1 was implemented using commercially available IC and LSI chips. This choice was a natural consequence of the fact

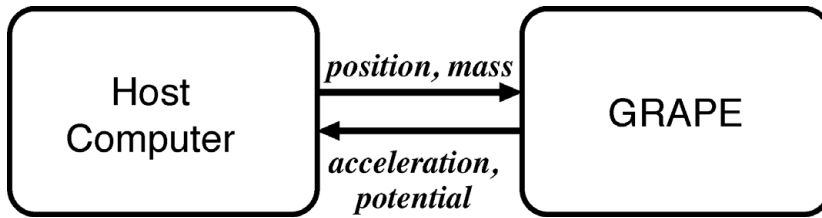


Figure 1. Basic structure of a GRAPE system.

that project members lacked both money and experience to design custom LSI chips. In fact, none of the original design and development team of GRAPE-1 had the knowledge of electronic circuits more than what was learned in basic undergraduate courses for physics students.

For GRAPE-1, an unusually short word format was used, to make the hardware as simple as possible. Except for the first subtraction of the position vectors (16-bit fixed point) and final accumulation of the force (48-bit fixed point), all operations are done in 8-bit logarithmic format, in which 3 bits are used for the “fractional” part. This choice simplified the hardware significantly. The use of extremely short word format in GRAPE-1 was based on the detailed theoretical analysis of error propagation and numerical experiment (Makino *et al.* 1990).

GRAPE-2 was similar to GRAPE-1A, but with much higher numerical accuracy. In order to achieve higher accuracy, commercial LSI chips for floating-point arithmetic operations such as TI SN74ACT8847 and Analog Devices ADSP3201/3202 were used. The pipeline of GRAPE-2 processes the three components of the interaction sequentially. So it accumulates one interaction in every three clock cycles. This approach was adopted to reduce the circuit size. Its speed was around 40 Mflops, but it is still much faster than workstations or minicomputers at that time.

GRAPE-3 was the first GRAPE computer with custom LSI chip. The number format was the combination of the fixed point and logarithmic format similar to what were used in GRAPE-1. The chip was fabricated using $1\mu\text{m}$ design rule by National Semiconductor. The number of transistors on a chip was 110K. The chip operated at 20MHz clock speed, offering a speed of about 0.8 Gflops. Printed-circuit board with 8 chips were mass-produced, for a speed of 6.4 Gflops per board. Thus, GRAPE-3 was also the first GRAPE computer to integrate multiple pipelines into a system. Also, GRAPE-3 was the first GRAPE computer to be manufactured and sold by a commercial company. Nearly 100 copies of GRAPE-3 have been sold to more than 30 institutes (more than 20 outside Japan).

With GRAPE-4, a high-accuracy pipeline was integrated into one chip. This chip calculates the first time derivative of the force, so that a fourth-order Hermite scheme (Makino & Aarseth 1992) can be used. Here, again, a serialized pipeline similar to that of GRAPE-2 was used. The chip was fabricated using $1\mu\text{m}$ design rule by LSI Logic. Total transistor count was about 400K.

The completed GRAPE-4 system consisted of 1728 pipeline chips (36 PCB boards each with 48 pipeline chips). It operated on 32 MHz clock, delivering the speed of 1.1 Tflops. Technical details of machines from GRAPE-1 through GRAPE-4 can be found in our book (Makino & Taiji 1998) and reference therein.

GRAPE-5 (Kawai *et al.* 2000) was an improvement over GRAPE-3. It integrated two full pipelines which operate on 80 MHz clock. Thus, a single GRAPE-5 chip offered a speed 8 times more than that of the GRAPE-3 chip, or the same speed as that of an

Table 1. History of GRAPE project

GRAPE-1	(89/4 — 89/10)	310 Mflops, low accuracy
GRAPE-2	(89/8 — 90/5)	50 Mflops, high accuracy(32bit/64bit)
GRAPE-1A	(90/4 — 90/10)	310 Mflops, low accuracy
GRAPE-3	(90/9 — 91/9)	18 Gflops, high accuracy
GRAPE-2A	(91/7 — 92/5)	230 Mflops, high accuracy
HARP-1	(92/7 — 93/3)	180 Mflops, high accuracy Hermite scheme
GRAPE-3A	(92/1 — 93/7)	8 Gflops/board some 80 copies are used all over the world
GRAPE-4	(92/7 — 95/7)	1 Tflops, high accuracy Some 10 copies of small machines
MD-GRAPE	(94/7 — 95/4)	1Gflops/chip, high accuracy programmable interaction
GRAPE-5	(96/4 — 99/8)	5Gflops/chip, low accuracy
GRAPE-6	(97/8 — 02/3)	64 Tflops, high accuracy

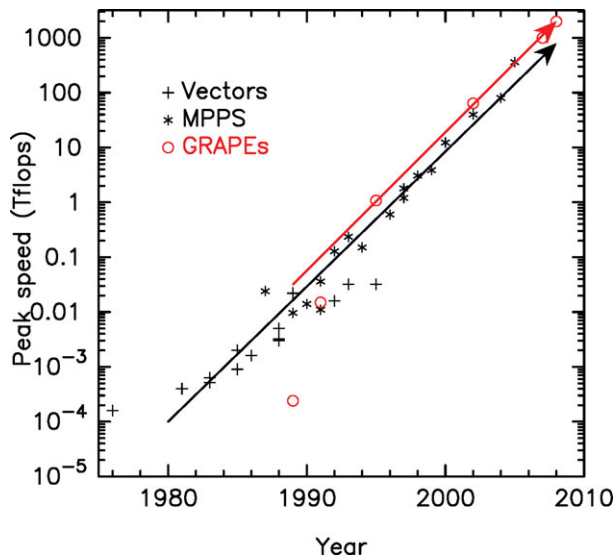


Figure 2. The evolution of GRAPE and general-purpose parallel computers. The peak speed is plotted against the year of delivery. Open circles, crosses and stars denote GRAPEs, vector processors, and parallel processors, respectively.

8-chip GRAPE-3 board. GRAPE-5 was awarded the 1999 Gordon Bell Prize for price-performance. The GRAPE-5 chip was fabricated with $0.35\mu\text{m}$ design rule by NEC.

Table 1 summarizes the history of GRAPE project. Figure 2 shows the evolution of GRAPE systems and general-purpose parallel computers. One can see that evolution of GRAPE is faster than that of general-purpose computers.

The GRAPE-6 was essentially a scaled-up version of GRAPE-4(Makino *et al.* 1997), with the peak speed of around 64 Tflops. The peak speed of a single pipeline chip was 31 Gflops. In comparison, GRAPE-4 consists of 1728 pipeline chips, each with 600 Mflops. The increase of a factor of 50 in speed was achieved by integrating six pipelines into one chip (GRAPE-4 chip has one pipeline which needs three cycles to calculate the force from one particle) and using 3 times higher clock frequency. The advance of the device technology (from $1\mu\text{m}$ to $0.25\mu\text{m}$) made these improvements possible. Figure 3 shows the processor chip delivered in early 1999. The six pipeline units are visible.

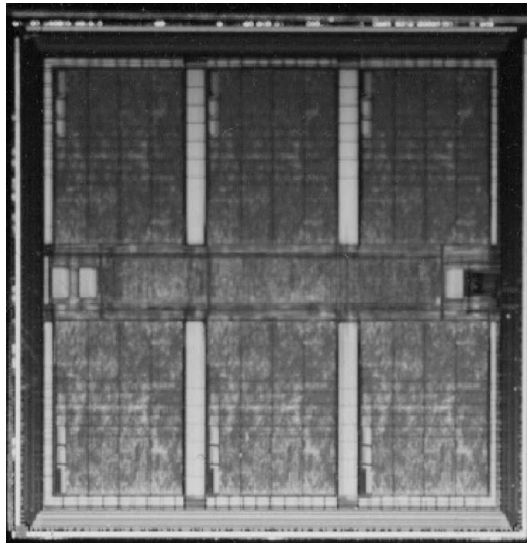


Figure 3. The GRAPE-6 processor chip.

The completed GRAPE-6 system consisted of 64 processor boards, grouped into 4 clusters with 16 boards each. Within a cluster, 16 boards are organized in a 4 by 4 matrix, with 4 host computers. They are organized so that the effective communication speed is proportional to the number of host computers. In a simple configuration, the effective communication speed becomes independent of the number of host computers. The details of the network used in GRAPE-6 is in Makino *et al.* (2003).

3. LSI economics and GRAPE

GRAPE has achieved the cost performance much better than that of general-purpose computers. One reason for this success is simply that with GRAPE architecture one can use practically all transistors for arithmetic units, without being limited by the memory wall problem. Another reason is the fact that arithmetic units can be optimized to their specific uses in the pipeline. For example, in the case of GRAPE-6, the subtraction of two positions is performed in 64-bit fixed point format, not in floating-point format. Final accumulation is also done in fixed point. In addition, most of the arithmetic operations to calculate the pairwise interactions are done in single precision. These optimizations made it possible to pack more than 300 arithmetic units into a single chip with less than 10M transistors. The first microprocessor with fully-pipelined double-precision floating-point unit, Intel 80860, required 1.2M transistors for two (actually one and half) operations. Thus, the number of transistors per arithmetic unit of GRAPE is smaller by more than a factor of 10. When compared with more recent processors, the difference becomes even larger. The Fermi processor from NVIDIA integrates 512 arithmetic units (adder and multiplier) with 3G transistors. Thus, it is five times less efficient than Intel 80860, and nearly 100 times less efficient than GRAPE-6.

However, there is another economical factor. As the silicon semiconductor technology advances, the initial cost for the design and fabrication of custom chips increases. In 1990, the initial cost for a custom chip was around 100K USD. By the end of the 1990s, it has become higher than 1M USD. By 2010, the initial cost of a custom chip is around

10M USD. Thus, it has become difficult to get a budget large enough to make a custom chip, which has rather limited range of applications.

There are several possible solutions. One is to reduce the initial cost by using FPGA (Field-Programmable Gate Array) chips. An FPGA chip consists of a number of “programmable” logic blocks (LBs) and also “programmable” interconnections. A LB is essentially a small lookup table with multiple inputs, augmented with one flip-flop and sometimes full-adder or more additional circuits. The lookup table can express any combinatorial logic for input data, and with flip-flop it can be part of a sequential logic. Interconnection network is used to make larger and more complex logic, by connecting LBs. The design of recent FPGA chips has become much more complex, with large functional units like memory blocks and multiplier (typically 18×18 bits) blocks.

Unfortunately, because of the need for the programmability, the size of the circuit that can fit into an FPGA chip is much smaller than that for a custom LSI, and the speed of the circuit is also slower. In order to be competitive, it is necessary to use much shorter word length. GRAPE architecture with reduced accuracy is thus an ideal target for FPGA-based approach. Several successful approaches have been reported (Hamada *et al.* 1999, Kawai & Fukushima 2006).

4. GRAPE-DR

Another solution for the problem of the high initial cost is to widen the application range by some way to justify the high cost. With GRAPE-DR project Makino *et al.* (2007), we followed this approach.

With GRAPE-DR, the hardwired pipeline processor of previous GRAPE systems were replaced by a collection of simple SIMD programmable processors. The internal network and external memory interface were designed so that it could emulate GRAPE processors efficiently and could be used for several other important applications, including the multiplication of dense matrices.

GRAPE-DR is an acronym of “Greatly Reduced Array of Processor Elements with Data Reduction”. The last part, “Data Reduction”, means that it has an on-chip tree network which can do various reduction operations such as summation, max/min and logical and/or.

The GRAPE-DR project was started in FY 2004, and finished in FY 2008. The GRAPE-DR processor chip consists of 512 simple processors, which can operate at the clock cycle of 500MHz, for 512 Gflops of single precision peak performance (256 Gflops double precision). It was fabricated with TSMC 90nm process and the size is around 300mm^2 . The peak power consumption is around 60W. The GRAPE-DR processor board (figure 4) houses 4 GRAPE-DR chips, each with its own local DRAM chips. It communicates with the host computer through Gen1 16-lane PCI-Express interface.

This card gives the theoretical peak performance of 819 Gflops (in double precision) at the clock speed of 400 MHz. The actual performance numbers are 640 Gflops for matrix-matrix multiplication, 430 Gflops for LU-decomposition, and 500 Gflops for direct N -body simulation with individual timesteps (figure 5). These numbers are typically a factor of two or more better than the best performance number so far reported with GPGPUs.

In the case of parallel LU decomposition, the measured performance was 24 Tflops on 64-board, 64-node system. The average power consumption of this system during the calculation was 29KW, and thus performance per Watt is 815 Mflops/W. This number is listed as No. 1 in the Little Green 500 list of June 2010. Thus, from a technical point of

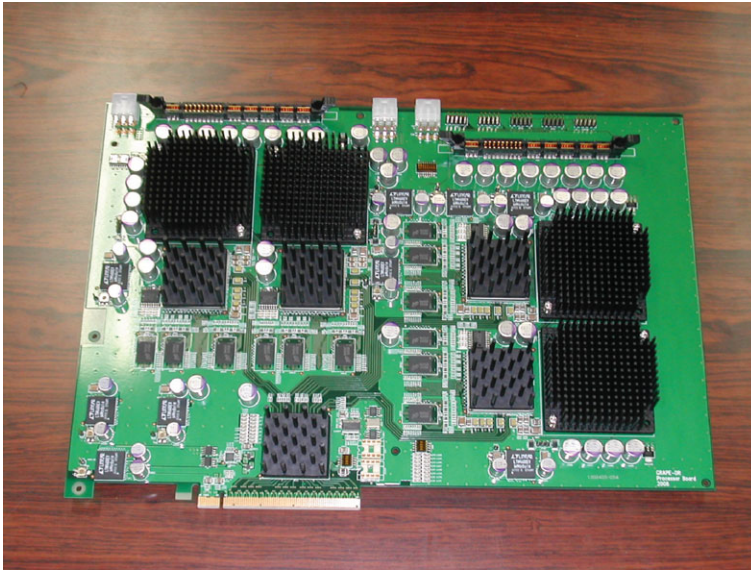


Figure 4. The GRAPE-DR processor board.

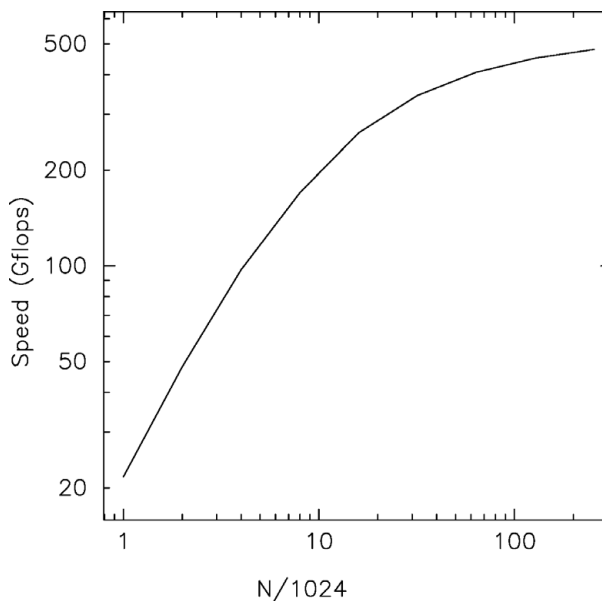


Figure 5. The performance of individual-timestep scheme on single-card GRAPE-DR in Gflops, plotted as a function of the number of particles.

view, we believe the GRAPE-DR project is highly successful, in making multi-purpose computers with highest single-card performance and highest performance-per-watt.

Whether or not the approach like GRAPE-DR will be competitive with other approaches, in particular GPGPUs, is at the time of writing rather unclear. The reason is simply that the advantage over GPGPUs is not quite enough, primarily because of the low production cost of GPGPUs. On the other hand, the transistor efficiency of general-purpose computers, and that of GPUs, have been decreasing for the last 20 years and probably will continue to do so for the next 10 years or so. GRAPE-DR can retain its



Figure 6. The GRAPE-DR cluster.

efficiency when it is implemented with more advanced semiconductor technology, since, as in the case of GRAPE, one can use the increased number of transistors to increase the number of processor elements. Thus, it might remain competitive.

5. Future directions

In hindsight, 1990s was a very good period for the development of special-purpose architecture such as GRAPE, because of two reasons. First, the semiconductor technology reached to the point where many floating-point arithmetic units can be integrated into a chip. Second, the initial design cost of a chip was still within the reach of fairly small research projects in basic science.

By now, semiconductor technology reached to the point that one could integrate thousands of arithmetic units into a chip. On the other hand, the initial design cost of a chip has become too high.

The use of FPGAs and the GRAPE-DR approach are two examples of the way to tackle the problem of increasing initial cost. However, unless one can keep increasing the budget, GRAPE-DR approach is not viable, simply because it still means exponential increase in the initial, and therefore total, cost of the project.

On the other hand, such increase in the budget might not be impossible, since the field of computational science as a whole is becoming more and more important. Even though a supercomputer is expensive, it is still much less expensive compared to, for example, particle accelerators or space telescopes. Of course, computer simulation cannot replace the real experiments of observations, but computer simulations have become essential in many fields science and technology.

In addition, there are several technologies available in between FPGAs and custom chips. One is what is called “structured ASIC”. It requires customization of typically just one metal layer, resulting in large reduction in the initial cost. The number of gates one can fit into the given silicon area falls between those of FPGAs and custom chips. We are currently working on a new fully-pipelined system, based on this structured ASIC.

The price of the chip is not very low, but in the current plan it gives extremely good performance for very low energy consumption.

References

- Barnes, J. & Hut, P. 1986, *Nature*, 324, 446
- Greengard, L. & Rokhlin, V. 1987, *Journal of Computational Physics*, 73, 325
- Hamada, T., Fukushige, T., Kawai, A., & Makino, J. 1999, PROGRAPE-1: A Programmable, Multi-Purpose Computer for Many-Body Simulations, submitted to *PASJ*
- Ito, T., Makino, J., Ebisuzaki, T., & Sugimoto, D. 1990, *Computer Physics Communications*, 60, 187
- Kawai, A. & Fukushige, T. 2006, \$158/GFLOP Astrophysical N-Body Simulation with a Reconfigurable Add-in Card and a Hierarchical Tree Algorithm
- Kawai, A., Fukushige, T., Makino, J., & Taiji, M. 2000, *PASJ*, 52, 659
- Makino, J. & Aarseth, S. J. 1992, *PASJ*, 44, 141
- Makino, J., Hiraki, K., & Inaba, M. 2007, in *Proceedings of SC07*, ACM, (Online)
- Makino, J., Ito, T., & Ebisuzaki, T. 1990, *PASJ*, 42, 717
- Makino, J., Fukushige, T., Koga, M., & Namura, K., T. 2003, *PASJ*, 55, 1163
- Makino, J. & Taiji, M. 1998, *Scientific Simulations with Special-Purpose Computers — The GRAPE Systems* (Chichester: John Wiley and Sons)
- Makino, J., Taiji, M., Ebisuzaki, T., & Sugimoto, D. 1997, *ApJ*, 480, 432
- Makino, J., Fukushige, T., Koga, M., & Namura, K., T. 2003, *PASJ*, 55, 1163