UDC 52-17: 681.3.06

VPP300 Series in National Astronomical Observatory

●Takashi Ito

(Manuscript received May 10, 1997)

In astronomy we cannot perform any experiments of realistic time scale and spatial scale because those of the universe are too large for human being. That is the reason why we make very much of the numerical experiments by large computers. In this paper, we introduce the supercomputer system VPP300/16R, VX/4R, VX/1R and many small computers around them which were installed in National Astronomical Observatory Mitaka, Tokyo in 1996. We also introduce the history, the way of administration, and some of the scientific results produced by the system for now. In this system main memory of 32 Gbyte and parallel computation up to 16PE is available, and it is expected to become a central platform for large scale numerical experiments to progress the astronomy worldwide.

1. Introduction

Astronomy is a science of observation. Even before the dawn of recorded history, we can be sure that humans looked up at the sky and marveled at the magnificence of the universe. However, because what they saw was so far away, marvel was all they could do.

Then came the modern age, in which experimental approaches were established for physics and chemistry that led to the discovery of many new facts and laws. However, throughout this period the study of the universe remained a science of observation. Of course, astronomers built telescopes and other observation apparatus, but they could not, for example, make a star or study the inside of black-holes.

Most astronomical phenomena evolve over extremely long periods, and our observations usually give the impression that the universe is static. Because of this apparent stasis, the only way astronomers have been able to explain what they see is to study similar parts of the universe and then formulate a plausible evolution that could account for all the observations. Therefore, unlike physics and chemistry, where scientists can perform full-scale experiments, astronomy has progressed as a mostly passive science and the only way astronomers can control the boundary and initial conditions of phenomena is through numerical calculation using computers. Therefore, numerical experiments are far more important in astronomy than in other sciences.

2. Astronomy and Numerical Experiments

Recent remarkable progress in technologies has enabled astronomers to obtain a very large amount of accurate observational data. The progress is particularly beneficial for astronomers because they investigate a universe that extends perhaps infinitely. Astronomers obtain numerous new findings as observations become more accurate, and further new findings emerge while they are still trying to understand the previous findings.

Many scientific fields have almost reached the limit of qualitative discovery by observation because they are almost near the limit of measurement accuracy. This situation makes the progress of astronomy quite unique because accuracy in astronomy is increasing. The following are some examples of recent developments in astronomy:

The Hubble Space Telescope, which was launched into orbit to avoid the effects of atmospheric fluctuation,[1] an adaptive optical technique for telescopes that can dynamically adjust the surface geometry of an 8 m main mirror by as little as 10^{-6} m,[2] a large submillmeter-wave array project that aims at a resolution accuracy of 0.01 arcsec by positioning more than 50 antennas 10 m in diameter along a single straight line,[3] a wide-area telescope project to locate stars to an accuracy of 10^{-5} arcsec using a relative VLBI to measure master sources from all over the Galaxy, and a gravitational wave astronomy study to detect gravitational waves from supernova explosions and other sources by extending the baseline of a Fabry-Perot interferometer to 300 m and measuring spatial strain to a relative accuracy of 10^{-21}.[4] With these efforts, our understanding of the universe has dramatically changed.

Let's take an example from planetology. Until 20 years ago, the term "planet" always meant a planet in our own solar system. However, recent studies have revealed the existence of protoplanetary disks in star-forming regions of the galaxy, and have provided significant information to researchers investigating planet formation.[5),6] Moreover, other recent studies have discovered planets orbiting around several stars, and the number of such discoveries is sure to increase in the future.[7]

The theoretical study of planetary system formation has been criticized by classical positivists because it does not permit direct comparison with the results of observations. However, the benefits obtained from current, accurate observations have completely silenced the criticism. Although we cannot dynamically observe the solar system's evolution over its 4.6-billion-year history, we can study other planetary systems that are at different evolutional stages and thereby gain insights into the evolution of our own.

Because of the dramatic improvement in observational accuracy due to new technologies, larger and more accurate numerical experiments are necessary to explain our observations. When PC clock speeds increase from today's 100 MHz range to 200 MHz, scientists will request clock speeds of 500 MHz. Likewise, when disk capacities are commonly of the G (10^9) byte-order, our observation data will require T (10^{12}) byte-order disk capacities. Supply, it seems, will never catch up with demand as far as computer performance is concerned.

We do not use numerical experiments only to explain observations. An accurate numerical experiment can be a guide to efficient observation. For example, the large optical infrared telescope, the Subaru, now being constructed by the National Astronomical Observatory of Japan at the top of Hawaii island, will have a main mirror of more than 8 m in diameter to observe stars as far away as 15 billion light-years.[2] Because observation time with these very large telescopes is in such demand, it is important that before starting their observations astronomers clearly know what is to be observed and how it can be observed efficiently. Numerical experiments can support this preparatory stage, and efficient use of numerical experiments will greatly increase the observational efficiency for large telescopes.

We often refer to a computer system used for numerical experiments and data analysis as a "telescope for theoretical astronomy." Astronomers who rely on observatory studies still require a theoretical telescope with a performance as high as that of leading optical and radio telescopes; this has been especially true in recent years because of the large amount of accurate data that has been accumulated. Considering these situations, this paper introduces the VPP300/16R supercomputer installed at the National Astronomical Observatory this year, and also looks at its peripheral systems and application methods.

The supercomputer has just been introduced and has not yet provided us with much outstanding results; therefore, this paper introduces the results of several trial calculations and discusses the system's future prospects. The National Astronomical Observatory computer system can perform not only theoretical numerical experiments, but can also process and analyze the large amount of data obtained from the above mentioned telescopes. The international de facto standard software for astronomical image data analysis, IRAF, has been installed in the VPP system in cooperation with Fujitsu system engineers (see the cover photo). Using the combination of the system's high vector-performance and large main memory, IRAF will provide a superb data-analysis environment.

3. Reasons for introducing a parallel supercomputer

The National Astronomical Observatory of the Ministry of Education, Science, and Culture was established in 1988 to play a central role in Japanese astronomy by unifying the Tokyo Astronomical Observatory of the University of Tokyo, the 3rd Department of the Atmospheric Research Center, Nagoya University, and the Latitude Observation Center.

The first computer was introduced there in March 1966 and was an OKITAC 5090D. This was followed by a FACOM 230-58, UNIVAC 1100-80B, FACOM 380R, and FACOM M780/10S. The FACOM M780/10S was in use up to last year. The CPU-usage rate of the FACOM M780/10S in 1995 was already below 30%. Therefore, a few years before the start of the decline of large-scale computers, we started to look for a computer system for the National Astronomical Observatory for joint-use by universities.

There are many requirements for a computer system used as a theoretical telescope, and we decided that a large main memory capacity rather than calculation speed was our top requirement. A large storage area is critical for astronomical numerical experiments and data analyses because of the need to process large arrays. A small calculation can be managed by continuously running a workstation for months. However, a large numerical experiment with a lattice of $500 \times 500 \times 500$ or $200 \times 200 \times 200$ particles cannot be performed by a workstation with a main memory of around 100 megabytes. This is the reason we requested a system with a very large main memory.

The official invitation to tender announced on July 20, 1994 invited applicants to meet the following basic requirements: 1) A main memory capacity of 8.5 Gbytes or larger and an extended memory capacity of 12 Gbytes or larger, and 2) a single-processor performance of 1 GFLOPS or higher and an overall system performance for a $10{,}000 \times 10{,}000$ LINPACK of 15 GFLOPS or higher.

Although Japan/U.S. trade conflicts made it difficult to purchase a computer that was labeled as a "supercomputer," we insisted on these "super" specifications so we could obtain a large theoretical telescope for the development of astronomy.

4. System Outline

A Fujitsu supercomputer, the VPP300/16R, and its peripheral systems were introduced in January 1996, and the system's official joint use was started in April. The following paragraphs outline the current computer system at the National Astronomical Observatory (**Figs. 1** and **2**).

The supercomputer system consists of a VPP300/16R, three VX/4Rs, and a VX/1R. Each PE has a main memory of 2 Gbytes, so a numerical experiment can occupy up to 8 Gbytes of the VX/4R and up to 32 Gbytes of the VPP300.

Such large main memories require even larger auxiliary memory devices, for example, magnetic

Fig.1— Supercomputer system VPP300/16R, VX/4R, VX/1R installed in National Astronomical Observatory (NAO), Tokyo, Japan.

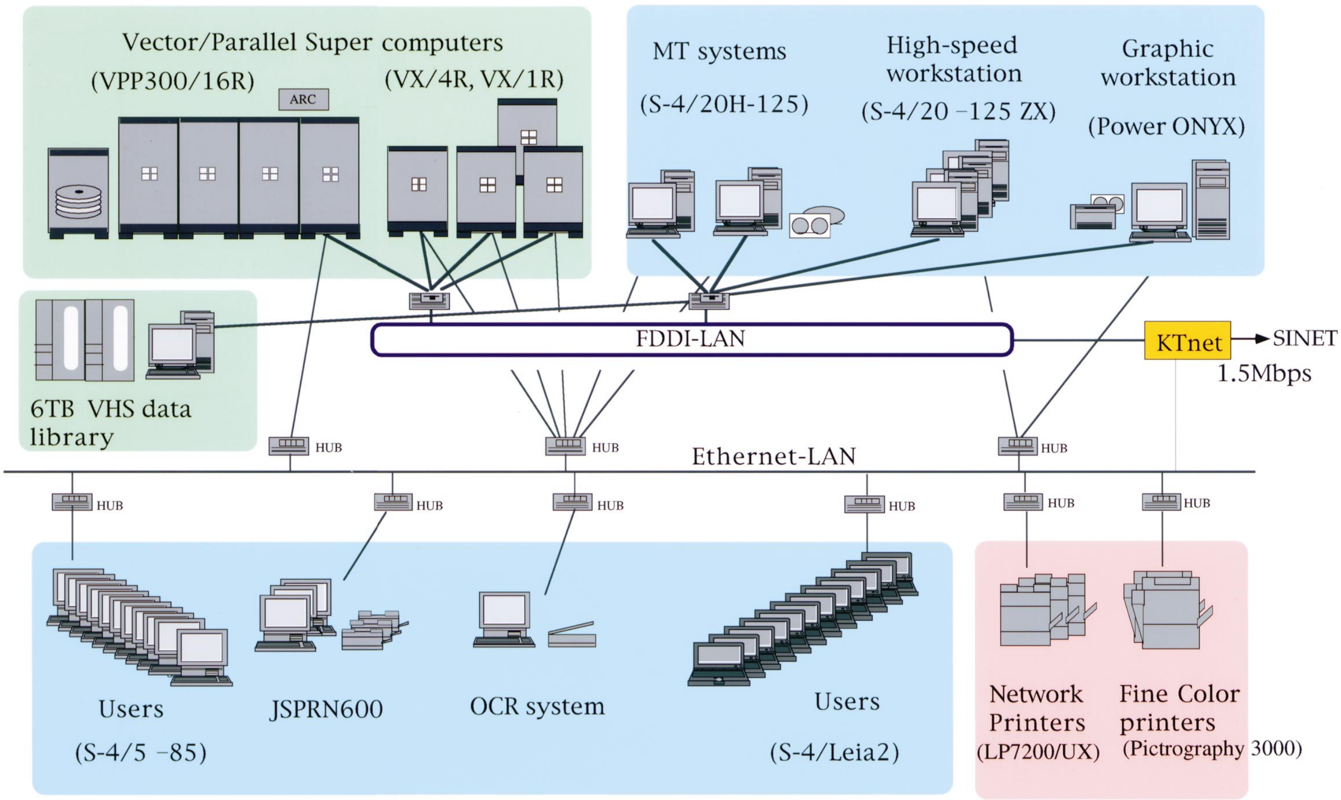

Fig.2— Configuration and distribution of supercomputer systems in NAO (October 1996). KTnet is the local area network of NAO, and SINET is the Internet backbone of NACSIS (NAtional Center for Science Information Systems).

disks. Therefore, magnetic disk devices with a total capacity of more than 200 Gbytes and a VHS tape library with a capacity of 6 Tbytes (ASACA) were added to the system. The supercomputer system, magnetic disk devices, and VHS tape library were installed in the underground computer room with dedicated power source and air-conditioning systems. Users can access the system through a 100 Mbyte/FDDI network.

Many research organizations usually stop their large computer system during nights and weekends (as we did with the previous system) to save electricity and allow time for maintenance. However, our new system is operated 24 hours a day so that users can do their work without interruption. (Such continuous operation is of great benefit to users, but it puts a heavy load on maintenance personnel.) This supercomputer runs jobs using a job control software product, Network Queuing System (NQS), to input a job from a remote workstation to a pipe queue. About 20 Fujitsu S Family workstations are provided as terminals to input jobs. Most of these terminals are in the computer terminal room in the astronomical observatory, and are running 24 hours a

Fig.3— Image processing workstation Power ONYX (SGI), and video editing system Turbo CUBE

day to work with the supercomputer system. Therefore, users can input a job any time they like.

In addition to the workstations, this system has various I/O devices, including three high-speed network printers (LP7200, RICOH), two color printers (Pictography 3000, Fuji Film), magnetic tape devices such as 8 mm cartridge tape and DAT drives, CD-ROM drives, and magneto-optical disk drives. Therefore, the system can satisfy almost every I/O need.

One of the most important features of our system is the image processing and editing system, in which the high-end workstation, Power ONYX (Reality Engine^2, Silicon Graphics), plays a central role (**Fig. 3**). We installed powerful image processing software, for example, Advanced Visualization System (AVS), and introduced an image editing device called Turbo CUBE to the ONYX so that the system can efficiently present scientific and technical computation results using advanced graphics. We sometimes invite the experts of the system's devices and software to hold small lecture meetings for users.

Some of the computation results that will be introduced later were prepared with the Power ONYX and AVS. In a study based on computation results, not only must the data be reliable, but we must also process it so that it can be presented in an easy-to-understand form. Therefore, the value of the supercomputer is greatly increased by these visualization systems.

Table 1. Large-scale numerical simulation study projects of the first half of 1996. The last two projects are developmental studies of a parallel-computation basic algorithm and are not directly related to astronomy.

- Impact processes on planetary surfaces and of planetary bodies
- Numerical simulation of relativistic extragalactic jets in a magnetic field
- Optimization of particle code on a parallel computer
- Numerical simulations of accretion flows
- Post-Newtonian simulation of coalescing binary neutron stars
- Coalescence of binary neutron stars and gravitational radiation
- Quasi-analytic approach to the problem of galaxy formation and analysis of growth of cosmological density fluctuations
- MHD simulation of flares and jets in astrophysics
- Realistic modeling of magnetic fields in the solar corona
- Parallel algorithms for extrapolation technique
- Numerical experiment of dynamics of rotating fluid as a model of polar regions of solar convection zone
- Theoretical determination of the solar wind and magnetic field
- Escape efficiency of impact ejecta from small bodies
- Flows around astronomical bodies
- Radiation hydrodynamical evolution in star formation
- Evolution of molecular abundance in protoplanetary disks
- Collapse and fragmentation of magnetized molecular clouds
- Gravitational instabilities on circumstellar disks
- 3-D mapping of source of solar oscillations
- Dissipation of protoplanetary disks
- Multidimensional simulation of turbulent nuclear flames in type Ia supernovae
- Stability analysis of general relativistic rotating stars
- Theoretical calculations of nucleosynthesis in stars
- The formation and evolution of dwarf galaxies
- Structure of the interstellar medium driven by supernova explosions
- Study of mass formula of nuclei and versatile nuclear reaction network
- Angular momentum transport and mass outflows in differentially rotating magnetized disks
- Cosmological simulation for the wide field galaxy and quasar survey
- A study of star and planetary system formation by radiation hydrodynamics
- Dynamical collapse of rotating, magnetized gas clouds
- Non-thermal emissions from shell-type supernova remnants
- Magnetodynamic simulation studies of active astronomical objects
- Thermal instability of the accretion disk in a close binary system
- Gravitational collapse of magnetized rotating cloud
- The efficient usages of parallel computers for the gravitational N-body problem
- Energy storage and liberation in solar corona
- Numerical study on quasi-stationary state of binary neutron star
- Simulation of galaxy formation
- Numerical studies on the problem of fueling for the AGN
- Hydrodynamics and magnetohydrodynamics in active astronomical objects
- Numerical simulations of formation and evolution of galaxy clusters
- MHD simulation of flares and jets in astrophysics
- Numerical simulation of the magnetic distortion on the solar surface and of solar high-energy events based on magnetic reconnection
- Study of star and planetary system formation by radiation hydrodynamics
- The evolution and formation of stars/disks with radiation hydrodynamics
- Numerical simulations of magnetohydrodynamic instabilities in galaxies and galaxy cluster
- Hydrodynamic simulation of coalescing binaries using SPH method
- Parallel algorithm towards 15 GFLOPS
- Evaluation of functionality of parallel processors

5. Administration

Unlike the large computer centers of other national universities, the National Astronomical Observatory traditionally has not charged users for the use of its large computers and has only asked them to fill out an application form. However, we have now established a procedure for allocating time for VPP300 and VX usage. The new procedure consists of usage application, application examination, and time allocation (similar systems have been in general use in astronomy, especially for allocating time for a telescope). The examiners appointed by the manager of the National Astronomical Observatory allocate time to applicants according to the importance of their study subjects and the outcomes of the applicants' work over the previous year. **Table 1** lists the 43 study projects of the first half of this year.

Our new system, from the terminal workstations to supercomputer, uses UNIX (UXP/V, Solaris, SunOS, and IRIX, etc.) as its operating system. UNIX is not suitable for data sharing by multiple investigators because it was developed by just a few investigators for their private use. In fact, the default function of UNIX cannot restrict the allocation of memory and CPU time to users' jobs. We therefore introduced the NQS to solve this problem. When a user inputs a job, a job class suitable for the job size is used to allocate memory and CPU time (**Tables 2**, **3**, and **4**). A function to periodically check the total PE time used by each study group is also essential for running the system according to the time allocation system. Therefore, a routine was incorporated into the NQS to check the accounting information for each user once a day and obtain statistical information about each group. Unlike conventional operation systems such as MSP, UNIX does not provide these functions as standard because it has not traditionally been concerned with job control. However, demands for UNIX job and group control tools will increase in the future. Because the system is too large to be maintained solely by the staff of the National Astronomical Observatory, three Fujitsu system engineers have been assigned to deal with problems and give users support. In addition, Fujitsu HPC Headquarters' special engineers perform on-site servicing for parallelization and vectorization programming in VPP Fortran twice a week.

Table 2. Job classes in the supercomputer of the National Astronomical Observatory, VPP300/16R. As of October 1, 1996, the number of acceptable Qs indicates the maximum number of jobs that can be input by the same study group simultaneously.

Queue name	Memory(M bytes)	Number of PEs	Multiplexity	Time(h)	Number of acceptable Qs
vpp4	7 616	4	2	8.0	2
vpp8	15 232	8	1	8.0	1
vpp15	28 560	15	1	8.0	1
vpp1	1 904	1	3	2.0	3
vpp4s	7 616	4	2	0.5	2
vpp8s	15 232	8	1	0.5	1
vpp15s	28 560	15	1	0.5	1

Table 3. Job classes in the supercomputer of the National Astronomical Observatory, VX/4R. As of October 1, 1996, there are three VX/4Rs, two of which are for open use. The following table shows the data for two VX/4Rs.

Queue name	Memory(M bytes)	Number of PEs	Multiplexity	Time(h)	Number of acceptable Qs
vx1	1 904	1	3	0.5	3
vx1s	944	1	1	8.0	3
vx2	1 888	2	1	2.0	3
vx3	2 832	3	1	2.0	3
vy1	1 520	1	3	0.5	3
vy4	6 080	4	1	8.0	3

Table 4. Job classes in the supercomputer of the National Astronomical Observatory, VX/1R (non-parallel computer), as of October 1, 1996.

Queue name	Memory(M bytes)	Multiplexity	Time(h)	Remarks:
a	64	6	0.5	
b	128	2	0.5	
c	256	1	0.5	
d	64	1	1.0	
e	128	1	1.0	Special application is required.
f	512	1	2.0	

6. Numerical Calculation Examples

We will need more time to study the effects of introducing the VPP system because only six months have passed since the start of full-scale joint-use (since April 1996). One reason why more time is required is that the vectorization and parallelization programs are difficult for users who are accustomed to sequential scalar computation on workstations. However, users are currently conducting studies with their own goals and guidelines, and some of these studies are close to providing attractive scientific outcomes. The following paragraphs introduce some numerical computation results obtained using parallelization, which most characterizes the VPP system.

6.1 Parallelized numerical solution for hydrodynamics

Condensation of interstellar gas into a star is often modeled using fluid approximation. Not only stars composed of gas, but also planets composed of rocks, such as the earth, behave as viscous fluids over an astronomically long time-scale. Therefore, many of the numerical experiments performed in astronomy are very similar to the numerical experiments performed in hydrodynamics. (Astronomers have always been major users of hydrodynamic supercomputing.)

Hydrodynamic numerical experiments on large matrices are most suitable for vector computers like supercomputers, and hydrodynamic numerical solutions have been found in various fields using vector computers. Modern supercomputers, in fact, have made it possible to numerically analyze the fundamental equations of hydrodynamics, and we could almost say that they were developed for hydrodynamic numerical solutions.

There are two methods of describing a fluid on a computer. These are the Euler method, which describes fluid motion using fixed coordinates and the Lagrange method, which describes a fluid using coordinates that move with the fluid). The hydrodynamic solution by the Euler method is popular. Many of the traditional numerical solutions for hydrodynamics are suitable for parallelization, and current calculation code can be paralellized with minor modifications.

The Euler method prepares a fixed lattice in space to determine the change in physical quantity at each point of the lattice over time. Even when a quadratic explicit method is used, determining a physical quantity x_i^n at step *n* and at lattice point *i* requires only three pieces of information; that is, the physical quantity at the point at step *n* and the physical quantities at the two adjacent points at the previous step. Mathematically, these three are denoted by x_{i-1}^{n-1} , x_i^{n-1} , and x_{i+1}^{n-1} .

Therefore, when a computation region is wide, it can be divided into several parts for parallel computation. Except for the boundary areas, each part can be computed independently. **Figure 4** shows a typical algorithm. This algorithm calculates the time series evolution of a two-dimensional fluid model. The computation space is divided into several large parts along the y axis for numerical fluid computation and time integration of various conservative quantities, and each part

```
Initial conditions (density, temperature, and pressure distribution) are given.
!XOCL PARALLEL REGION
Do: Time evolution loop
    !XOCL SPREAD DO
    DO: y direction                ←Parallelization
        DO:  x direction           ←Vectorization
            Numerical flux is computed.
        End do
    End do
    !XOCL END SPREAD DO
    !XOCL OVERLAPFIX (Numerical flux)
    DO: y direction                ←Parallelization
        DO:  x direction           ←Vectorization
            Various conservation quantities are time-integrated.
        End do
    End do
    !XOCL END SPREAD DO
    !XOCL OVERLAPFIX (Conservation quantity)
End do
!XOCL END PARALLEL
```

Fig.4— An example of parallelized algorithm for solving equations of motion for fluid dynamics. Space for computation are divided in large parts along y axis. To calculate interactions between each parts, it is necessary to make OVERLAPFIX in the source codes.

is parallelly computed by independent PEs. This method is much quicker than non-parallel computation with a single PE (**Fig. 5**). Using a large main memory enables us to use bigger lattices and greatly enhance accuracy. **Figure 6** shows an example of the initial evolution of a protoplanet disk that was calculated using this method.

When the most common type of supercomputer was the vector supercomputer, more than half of the applications were for finding numerical solutions to hydrodynamics problems. Astronomers are continually developing new parallelization and vectorization algorithms to improve computation speed and accuracy, and further development is expected.

6.2 Parallel Computation for Radiative Transfer Processes

Radiative energy transfer often plays an important role in astronomical structural formation and dynamic evolution. In the study of planet formation, the quantitative evaluation of radiative transfer processes may provide important information about, for example, the gravitational contraction of a molecular cloud core, the formation of a protostar core and surrounding disk, and the gravitational instability of a disk surrounding a star.[8] A radiative hydrodynamics theory that deals strictly with the interaction between a radiative field and surrounding material would involve high-dimensional (three-dimensional spatial, two-dimensional directional, and one-dimensional wavelength) equations and thus require a huge amount of computation. Although these kinds of numerical computations are performed based on certain assumptions, they are insufficiently accurate and may even give a qualitatively incorrect solution under certain boundary conditions.

To overcome this problem, radiative fluid dynamics computation algorithms suitable for parallelized computations are now being devel-

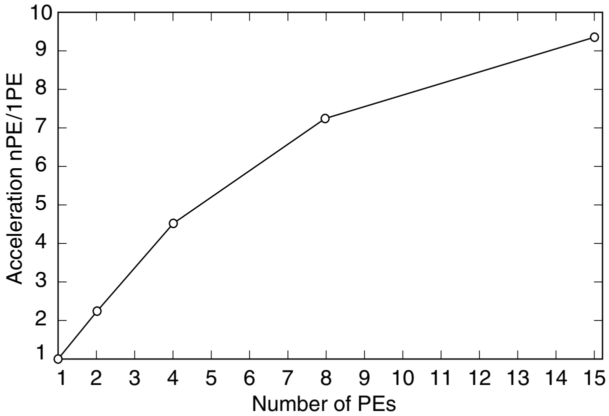

Fig.5— Improvement of computational performance using parallelized algorithm for fluid dynamics. Horizontal axis denotes the number of PE used, vertical axis denotes the ratio of computation time of parallel calculation to sequential calculation. Here we use the computational method described in Figure 4. Number of mesh is 256 × 256.Performance of parallelization seems to go down as the number of PE increases. This is perhaps because of the FFT routine in SSL II VPP dp_v2drcf. See also Figure 10.

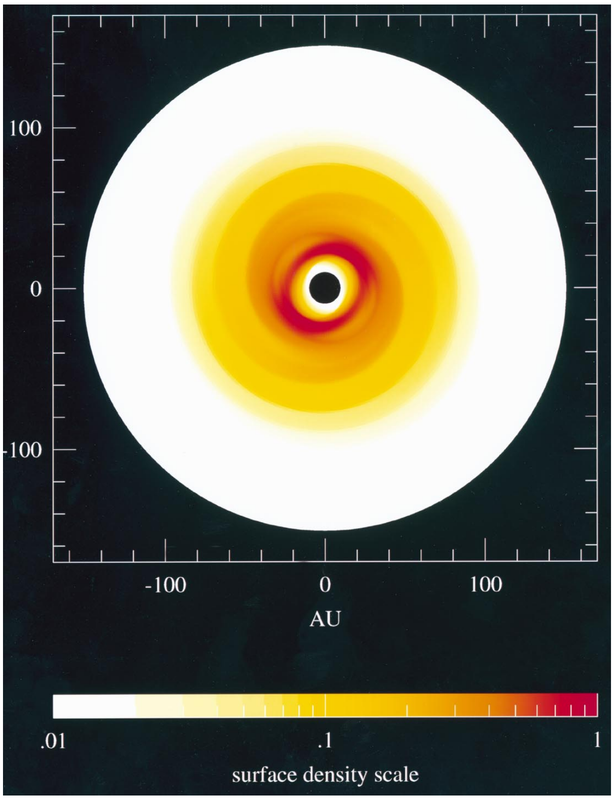

Fig.6— An example which shows the gravitational instability of planetary nebula calculated by the parallelized algorithm (Figure 4) on VPP300. Color bar shows the surface density in which red denotes high density and white means low density area. Unit of axes is Astronomical Unit, AU.

oped[9] which will enable us to accurately calculate radiative transfer in multidimensional space. One such algorithm is outlined in **Fig. 7**. This algorithm is divided into Processes A and B, and a simple estimation shows that Process B requires 100 times or more computation time than Process A. However, for the reason explained below, this might not be such a problem.

Process A solves a steady-state radiative transfer equation with the Short Characteristic Method, which divides space into a lattice of unit cubes and then determines the radiative intensity at each lattice point for all spatial directions and all wavelengths.[10] This method enables us to compute the radiative intensity for each spatial direction and each wavelength independently, which might greatly contribute to efficient parallelization in Process B.

Several preliminary numerical computations have met our expectations for high speed (**Fig. 8**). **Figure 9** shows the results of the preliminary computations. Figure 9 shows a radiative source at the center of a cube and the consequent radiative energy density per unit volume around the source.

```
The initial value of the Eddington factor is given.
!XOCL PARALLEL REGION
Do: time development loop
   Time integration of a radiative hydrodynamic equation ··· Process A
   Do: For each wavelength                                   Process B
      !XOCL SPREAD DO
      Do: For each direction              ←Parallelization
         For each lattice point           ←Vectorization
            Radiative intensity calculated (for each lattice point and direction).
         End do
      End do
      !XOCL END SPREAD DO
   End do
   Radiative intensity is integrated for each lattice point and
   direction to calculate the Eddington factor.
End do
!XOCL END PARALLEL
```

Fig.7— An example of parallelized algorithm for calculation of radiative transfer.It consists of two parts : numerical integration of radiative fluid dynamics using fixed Eddington factor (process A), and solving the equation of stationary radiative transfer under the temperature and density field obtained in process A (process B). Although process B requires more than 100 times of computational amounts than process A,it can be parallelized and vectorized quite effectively.

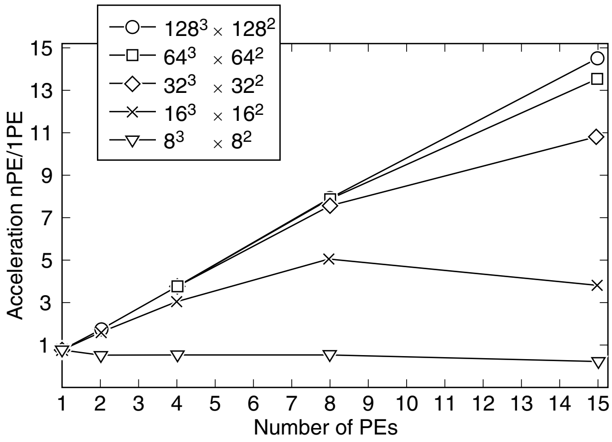

Fig.8— Improvement of computational performance using parallelized algorithm for radiative transfer. Horizontal axis denotes the number of PE used, vertical axis denotes the ratio of computation time of parallel calculation to sequential calculation. Here we use the computational method described in Figure 7. Number in the figure shows the number of mesh × division number of direction. In the case of large data (i.e. many meshes), it seems to be quite well parallelized.

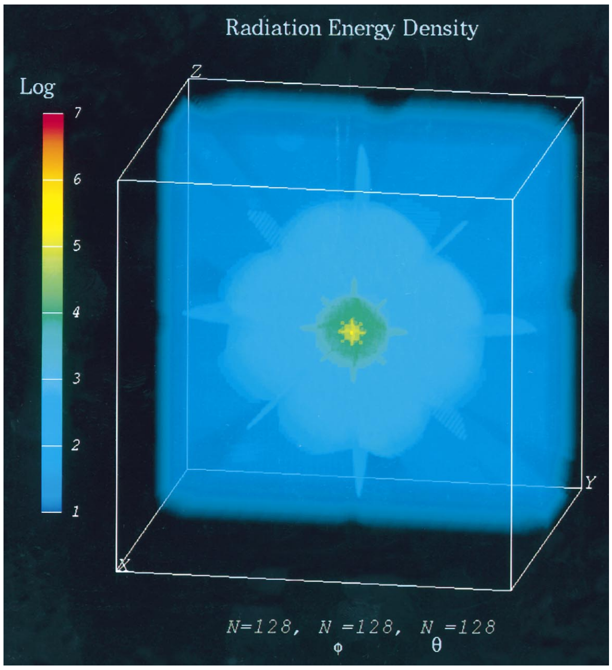

Fig.9— An example which shows the density distribution of radiative energy calculated by the parallelized algorithm (Figure 7) on VPP300. This figure was drawn by AVS on Power ONYX.

The data was obtained by running the algorithm on a VPP300 and then subjecting it to three-dimensional volume-rendering by AVS on the Power ONYX.

Contrary to the theoretical predictions, the result shows a non-spherical symmetry; this contradiction is due to insufficient spatial meshes. However, this result was obtained much quicker and shows a far higher accuracy than those obtained from conventional small-scale computations.

Calculating radiative transfer with realistic boundary conditions has been the most difficult problem in the investigation of star and planetary system formation; however, it is expected that further improvement of the algorithms and code will enable us to solve the problem.

6.3 Data amount and parallelization effect

The previous two computation results show that a relatively large amount of data is required to obtain the benefits offered by the VPP300 vector parallel supercomputer. This is more evident if we examine the results of some benchmark tests for data amount and parallelization efficiency. **Figure 10** shows the partial benchmark results of function dp_v2drcf of the SSL II VPP, which is a VPP version of the Fujitsu Scientific and Technological Computation Library, SSL II.[11] The figure clearly shows that increasing the number of PEs improves the computation speed only when the matrix is 1,024 × 1,024 or larger, and that increasing the number of PEs with matrixes 512 × 512 and smaller actually has a detrimental effect. One explanation for this is that when the data amount is small, the effect of parallelization may be offset by inter-PE communication bottlenecks, no matter how many PEs are used.

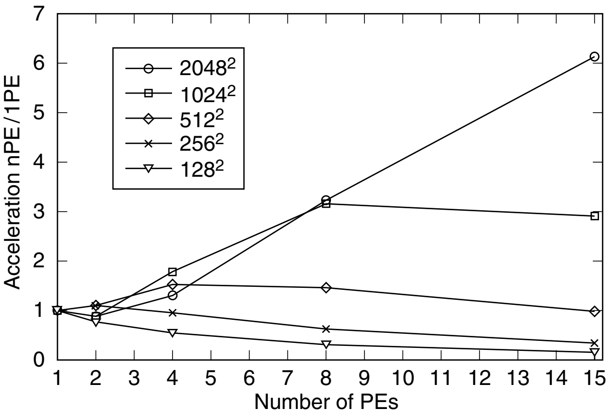

Fig.10— One of the benchmark results of FFT subroutine in SSL II VPP dp_v2drcf. Horizontal axis denotes the number of PE used, vertical axis denotes the relative computational time to sequential computation in the case of FFT of matrix-shaped two dimensional data on VPP300.

It has been generally understood that a large amount of data is required to realize the full potential of a large vector computer. This requirement also applies to parallelized computation, where the data amount (or array size) is a key factor in computation efficiency — which is why we requested a huge main memory for our new system. However, the VPP system is not always poor in dealing with small amounts of data.

Most of the personal computers and workstations currently available to investigators contain only one PE; however, because of the higher computation efficiency of parallel computation, computers with multiple processors such as the VPP system will become common in the near future. The VPP system can be used as a platform to develop numerical algorithms dedicated to parallel computation. In fact, some of the study projects that use the National Astronomical Observatory computers shown in Table 1 will develop algorithms for parallel computation. The following section introduces one of these projects, namely our efforts to speed up the numerical computation of a gravitational few-body problem through parallelization.

6.4 Solution of ordinary differential equations by parallelized extrapolation

The accurate determination of the rotation and revolution of planets, asteroids, and comets based on gravitational interactions is a major issue in several important questions in astronomy and planetology, for example, the origin, evolution, and stability of the solar system;[12] the internal structure and formation process of planets;[13),14] and the change in the earth's climate system and life evolution.[15)]

Gravitational interactions according to classical principles are not so difficult to determine. To determine them we need only calculate the acceleration due to gravitation between the bodies for all positions, and then solve the time evolution of the equations of motion.

However, these gravitational few-body problems (this term refers to gravitational interactions among a small number of bodies, and is used to distinguish from the frequently used term "gravitational many-body problem") have been considered unsuitable for vectorization and parallelization because of the small number of bodies they involve. For example, when dealing with the dynamical motions of our solar system, the number of bodies is only about 10, and the number falls to 9 when dealing only with the planets. Our studies have showed that vectorizing such a small number of components reduces computation speed.

Force function matrix computation is an important technique that is considered to be a parallel technique. However, actually parallelizing it may result in slower speeds because the effect of multiple PEs will be offset by inter-PE communication bottlenecks (Fig. 10). Fortunately, by changing the viewpoint slightly, the initial-value problem of the ordinary differential equation can be efficiently solved in parallel. A critical point is to understand that parallelization is performed not for particles, but for each stepsize . This idea is realized using an extrapolation method.

When an initial value problem of an ordinary differential equation is solved numerically, the virtual system on the computer gets closer to the actual physical system as the stepsize, h, approaches 0. A numerical method called the extrapolation method combines multiple solutions from the numerical integration of several finite stepsizes by using polynomial approximation, and then extrapolates a limit value of h to 0.[16)] This method, which was developed by Gragg, gives a far higher accuracy than conventional numerical integration methods.[17)] It has been used for accurate computations, for example, to calculate astronomical ephemerides, because certain iterative calculations provide maximum accuracy within a specific range.[18)] The extrapolation method involves much computation, and therefore has been considered unsuitable for calculating long-term evolutions such as those that determine whether a planetary system is stable or not. However, the computation for each h is independent, which enables multi-PE computation without inter-PE interference or the need for inter-PE communications, thereby ensuring effective and efficient parallelization of the extrapolation method.

Figure 11 outlines the algorithm, and **Fig. 12** shows the acceleration rates achieved by the parallelized extrapolation method for various numbers of PEs. When the number of orders for extrapolation is n, the stepsize for numerical integration is usually specified by the series 1/2, 1/4, 1/6, 1/8, ..., 1/2n, ... Because our experiment was performed with n=8, no acceleration can be expected even when eight or more PEs are used. Also, because the computation amount of each order is proportional to 2n, allocating the nth order to the nth PE results in unequal loads among PEs, which

```
The initial values of particles (positions and velocities) are given.
!XOCL PARALLEL REGION
Time development loop
  !XOCL SPREAD DO
  For each stepsize                  ←Parallelization
    For each dimension
      For each particle              ←Vectorization
        Force function for each particle and dimension is calculated.
      End do
    End do
    Particle position and velocity are calculated for each stepsize.
  End do
  !XOCL END SPREAD DO
  Values with a stepsize near to 0 are extrapolated by rational
  function extrapolation.
End do
!XOCL END PARALLEL
```

Fig.11— An example of parallelized algorithm for the gravitational N-body problem using extrapolation method. Parallelization is adapted for each step, not for each particle, so the performance for parallelization does not depend on the number of particles in this algorithm.

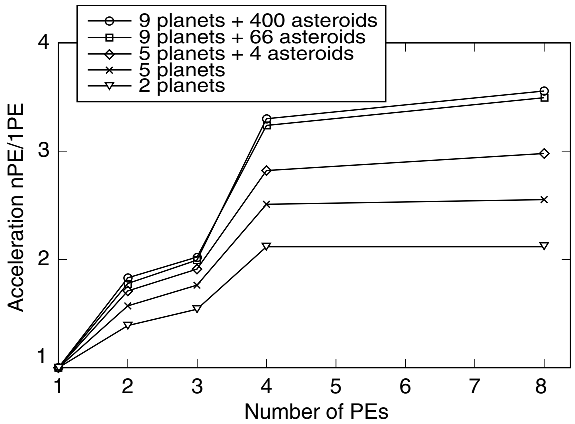

Fig.12— Improvement of computational performance using parallelized algorithm for extrapolation method. Horizontal axis denotes the number of PE used, vertical axis denotes the ratio of computation time of parallel calculation to sequential calculation. Since this is the result of n = 8 case, it is impossible to expect improvements over 8PE cases. In addition, because of the folding effect of each step, performance of 8PE case and 4PE case become nearly the same.

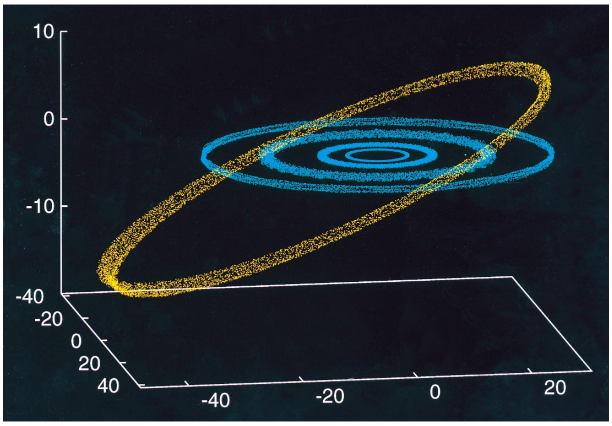

Fig.13— Result of numerical integration of outer planetary motion during ten million years using parallelized extrapolation method on VPP300. Each ring shows Jupiter, Saturn, Uranus, Neptune (cyan points), and Pluto (yellow points). Plotting interval is 400 thousand days. Unit of axes is Astronomical Unit (AU), and the coordinates are converted so that the ascending node of Pluto is always on x-axis. This figure was drawn by gnuplot on S-family workstation.

reduces the parallelization effect.

Therefore, we folded the part parallelized with the extrapolation method (combined 1 and 8, 2 and 7, etc.) so that each PE had an equal load. This brought an acceleration of about 3.5 times when four PEs were used. This performance is exceptionally good for the gravitational few-body problem, which has been thought to be basically unsuitable for parallelization. **Figure 13** shows a numerical integration of the results of extrapolation for the outer planets for about 10 million years. The figure shows, among other things, that Pluto maintains a peculiar resonance with Neptune.[18] It is expected that thoroughly verifying this motion will solve many questions about planetary accretion processes.

Because it is so time consuming, the extrapolation method has so far been used only to confirm computation accuracies in the long-term numerical computation of the gravitational few-body problem. However, with the increase in the number of parallel computers being used around the world, parallelized extrapolation may play a central role in the solution of the gravitational few-body problem. This study has shown that the effect of parallelization is maximum when the data amount is highest (9 planets and 400 asteroids) (Fig. 12).

7. Conclusion

Among the natural sciences, astronomy has been making exceptional progress. Our trial to provide a new approach to high-volume numerical computation is just getting under way with the aid of the high-performance computer system, VP300. The recent increase in Japanese astronomical research budgets is the highest so far probably because Japanese society has recognized the importance of natural science and researchers can now use high-performance computers for their studies.

However, supercomputers are affected by several negative factors which are largely independent of improvements in performance. Although current supercomputers can perform large-scale numerical experiments, they are inferior to workstations when the data amount is small and when data cannot be vectorized. Also, current supercomputers have to rely on FORTRAN (which has almost been replaced by C and Pascal in universities) to implement parallel computation. Therefore, it is natural for many young researchers and university students to choose research subjects that can be solved with their own personal computers or workstations.

Recently, special-purpose computers have become popular in Japanese astronomy, and projects to create high-performance special-purpose computers are under way or have been completed for the numerical computations described in this paper 19) and 20). Considering these facts, supercomputers will be rare in the 21 century unless their costs can be dramatically reduced. Japan/U.S. trade friction does not seem to be settled, and the procedure to purchase a large computer will become more complex. Mass-produced, inexpensive PCs supported by a large number of consumers will leave supercomputers behind in scalar computation. The necessity of supercomputers in natural science must therefore be carefully examined.

However, it is evident that current natural science studies cannot advance without computers, and that some astronomical (and other) fields can be studied only with a large-scale vector parallel computer. We expect that computer makers, including Fujitsu, will examine what is necessary for the development of the natural sciences without paying too much attention to the prefix "super."

Despite various problems and unexpected troubles, our system has been in full-scale operation since October 1996. We will have to wait and see whether this system will provide valuable scientific results or just disappear like the dinosaurs.

Acknowledgments

The author is deeply thankful to Dr. Nobuhiro Kikuchi of the National Astronomical Observatory, Dr. Taishi Nakamoto of the University of Tsukuba, and Dr. Masaru Hamabe of the Faculty of Science, University of Tokyo for providing me with some of the computation results and figures presented in this paper.

References

1) http: //www.stsci.edu/public.html.
2) http: //www.subaru.nao.ac.jp/.
3) http: //www.nro.nao.ac.jp/LMSA/lmsa_j-euc.html.
4) http: //www.tamago.mtk.nao.ac.jp/.
5) Beckwith, S. V. W. and Sargent, A. I: The occurrence and properties of disks around young stars, pp.521-542. The University of Arizona Press, Tucson & London, 1993.
6) Sargent, A. I. and Beckwith, S. V. W.: The detection and study of pre-planetary disks. Astrophysics and Space Science, 212, pp.181-189 (1994).
7) Walker, G. A. H. A solar system next door, Nature, 382, pp.23-24 (1996).
8) Nakamoto, T. and Nakagawa, Y.: Formation, early evolution, and gravitational stability of protoplanetary disks. Astrophys. J. 421, pp.640-650 (1995).
9) Nakamoto, T.: Radiation hydrodynamics for star and protoplanetary disk formation. In Tomisaka, K, editor, Numerical Astrophysics using Supercomputers, pp.11-13. National Astronomical Observatory, Mitaka, Tokyo, Japan, 1996.
10) Stone, J, M., Mihalas, D., and Norman, M.L.: ZEUS-2D: a radiation magnetohydrodynamics code for astrophysical flows in two space dimensions. III, the radiation hydrodynamic algorithms and tests. Astrophys. J. Suppl., 80, pp.819-845 (1992).
11) http: //www.cc.nao.ac.jp/whatsnew/bench-k.txt.
12) Tremaine, S.; Is the solar system stable? 1995. CITA-95-3, in Proceedings of Rosseland Centenary Symposium of Astrophysics Oslo, June 16-17, 1994.
13) Ito, T., Masuda, K., Hamano, Y., and Matsui, T.: Climate friction: a possible cause for secular drift of the earth's obliquity. J. Geophys. Res., 100, pp.15147-15161 (1995).
14) Ito, T. and Hamano, Y.: Evolution of the earth's obliquity and the role of core-mantle coupling.

In Yukutake, T., editor, The Earth's Central Part: Its Structure and Dynamics, pp.301-318. Terra Scientific Publ., Tokyo, 1995.

15) Ito, T., Kumazawa, M., Hamano, Y., Matsui, T., and Masuda, K.: Long term evolution of the solar insolation variation over 4 Ga. Proc. Jpn. Acad., Ser. B, 69, pp.233-237 (1993).

16) Hairer, E., Nosett, S. P., and Wanner, G.: Solving Ordinary Differential Equation I. Nonstiff Problems (second revised edition). Springer-Verlag, New York, 1991.

17) Gragg, W. B.; On extrapolation algorithms for ordinary initial value problems. SIAM J. Numer, Anal., 2, pp.384-403 (1995).

18) Kinoshita, H. and Nakai, H.: The motion of Pluto over the age of the solar system. In Ferraz-Mello, S. et al, editor, Dynamics, ephemerides and astrometry in the solar system, pp.61-70. Kluwer Academic publishers, Dordrecht, 1995.

19) Yokono, Y., Ogasawara, R, Takeuchi, T., Inutsuka, S., Miyama S. M., and Chikada, Y.,: Development of special-purpose computer for cosmic hydrodynamics by SPH. In Tomisaka, K, editor, Numerical Astrophysics using Supercomputers, pp.36-38. National Astronomical Observatory, Mitaka, Tokyo, Japan, 1996.

20) Sugimoto, D., Chikada, Y., Makino, J., Ito, T., Ebisuzaki, T., and Umemura, M.: A special-purpose computer for gravitation many-body problem. Nature, 345, pp.33-35 (1990).

Takashi Ito completed his studies at the graduate school of the Institute of Earth and Planetary Physics, Faculty of Science, the University of Tokyo, Japan in 1995.
He joined the Astronomical Data Analysis Center, National Astronomical Observatory, Mitaka, Tokyo in 1995, and has been engaged in research and development of planetary physics, especially numerical calculations on the stability of planetary orbits. He is a member of the Japan Society for Planetary Sciences and the Astronomical Society of Japan.