High performance computing beyond the peta scale in Japan

To cite this article: Toichi Sakata 2007 J. Phys.: Conf. Ser. 78 012059

View the article online for updates and enhancements.

Related content
- Biomolecular simulations on petascale: promises and challenges
  Pratul K Agarwal and Sadaf R Alam
- World's fastest supercomputer opens up to users
  Ling Xin
- Funding boost for Santos Dumont supercomputer
  Cássio Leite Vieira
High Performance Computing beyond the Peta Scale in Japan

Toichi Sakata
Executive Director, RIKEN,
2-1, Hirosawa, Wako, Saitama 351-0198, Japan

E-mail: tsakata@riken.jp

Abstract. Since last year, with funding from the Japanese government, RIKEN has been conducting a project to build a next generation supercomputer, which will be capable of carrying out peta or higher scale simulations and will be in full operation by the end of FY2011. Its main objectives are to keep leading edge hardware technology inside Japan, and also to stimulate the use of simulations in R&D in academia and industry in order to investigate sophisticated natural phenomena in ways which are only possible with such large scale simulations, for example, designing novel functional materials to create future industrial technologies. To promote this kind of use of the supercomputer, nano-science and life sciences have been selected as grand challenges, in which specially designed supercomputer codes are to be developed. This paper presents the key features of the project.

Key Word: Supercomputer, Peta Scale, National Project, Nano-science, Life Sciences, Industrial Usage

1. Introduction

Remarkable successes in predictions of climate change by the Earth Simulator have generated increased attention for the role of computational science as a predictive scientific approach, rather than a conventional descriptive one, used to explain experimental results. Computational science has been recognized as a third scientific method after theory and experiment, and has already become reliable enough to predict phenomena in many of research areas. Applications to industrial research have become a hot topic in the search for innovative production technologies.

On the other hand, the costs of developing high performance computers are increasing quite rapidly and therefore research and development into new, efficient and capable computers must be maintained to keep a technological leading edge. This is definitely evident for supercomputers. Furthermore, it is clear that the technologies of LSI devices and interconnections, employed in supercomputers, have always been leading and accelerating their improvements themselves. Considering this situation, the Government of Japan has decided to develop a next generation supercomputer as a national project in 2006, expecting that the hardware technologies developed in the project would expand them into commercial products like information devices. RIKEN has been conducting the project since 2006 with funding from the Japanese Government, and has already received positive attention from high performance computing communities worldwide.

In this article, basic policy on supercomputing in Japan and a brief introduction to the project are given.
2. Policy on Supercomputing in Japan
The Council for Science and Technology Policy (CSTP) is one of the four councils of important policies of Cabinet Office. The CSTP is comprised by Prime Minister, relevant Ministers, and distinguished experts. **Figure 1 briefly shows the Science and Technology Administration in Japan.** CSTP was established in January, 2001 inside the Cabinet Office as one of the governmental top councils based on the Law for Establishing Cabinet Office. Therefore, the future direction of science and technology in Japan is basically discussed at CSTP.

![Figure 1. Science and Technology Administration in Japan](image)

Every five years, CSTP proposes a “Science and Technology (S&T) Basic Plan” to be authorized by a Cabinet Meeting. This is considered to be essential to keep Japan a top player in a competitive environment. In the second S&T Basic Plan, covering the period from 2001 to 2005, the following four fields were chosen as areas for prioritized governmental investment: life sciences, information and telecommunications, environmental sciences, and nanotechnology/materials. In March last year, the Japanese Government announced the third S&T Basic Plan covering the period 2006 to 2010. **Figure 2 shows the outline of this plan.** This emphasized again the importance of the four research fields, saying “the core strategies of the third plan are: development of world-class researchers who can produce excellent research findings, creation of a competitive environment, promotion of science, and creation of persistent innovations through strategic investment; and removal of systematic or operational obstacles to return the R&D benefits back to society; and removal of systematic or operational obstacles to return the R&D benefits back to society” [1]. In the third plan, particular emphasis is being placed on promoting the development and shared use of advanced large-scale public research facilities. Specifically, it has been pointed out that the national government should take responsibility for development, and should promote the shared use of advanced, large-scale public research facilities such as next generation supercomputers and next generation synchrotron radiation sources such as XFEL. The next generation supercomputer, X-ray free electron laser, ocean and earth exploration system, fast breeder reactor (FBR) cycle technology, and space transport system have been chosen as Key Technologies of National Importance. Then, the Japanese government has entrusted RIKEN with the former two projects, which were launched last year.
Following this decision by the Japanese Government, the Office of Supercomputer Development Planning (OSDP) was established at the Research Promotion Bureau of MEXT about two years ago. This is an indication that MEXT is taking full responsibility for the next generation supercomputer project. Furthermore, among those innovative technologies discussed by the Japanese Government, the project is becoming the center of public attention, since the usefulness of simulations has been recognized recently in many pure and applied research fields as a third scientific method, after theory and experiment, and also from industry’s point of view.

After the establishment of OSDP in MEXT, the Next Generation Supercomputer R&D Center (NSC) was set up in RIKEN in January, last year, which is the operation body for the project. Then, taking the importance of the project into consideration, Dr. Noyori, the president of RIKEN himself has been taking the position of the NSC director. Now, the author is in the position of the deputy director of NSC and Dr. Watanabe is the project leader.

During the project, it will be important to have such occasional review meetings between OSDP and NSC in order to build public confidence in this supercomputer project.

3. Previous Supercomputer Projects in Japan
To date, there have been three major leading supercomputer projects in Japan. These are the Numerical Wind Tunnel, developed in 1993 by the National Aerospace Laboratory of Japan, CP-PACS, developed by the University of Tsukuba in 1996, and the Earth Simulator, developed by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) in 2002. All three recorded top performance at the TOP500 contests for three, one and three years, after they were built, with peak performances of 280, 614 GFLOPS, and 41 TFLOPS, respectively. After completing these projects, computers based on the device technology newly developed there were commercialized by Fujitsu, Hitachi and NEC as the VPP500, SR2201 and SX-6/SX-7, and have been used at national institutes and universities worldwide in various research fields. The next generation supercomputer project is not just their successor. That is, those supercomputers were developed just for one time, but the Japanese Government’s policy intention is for the project to continue, with development of successive supercomputers. Therefore, future ideas for supercomputers will also be explored through the present project.
Following the agreement of the Kyoto Protocol in 1997, the Japanese government has been carrying out many of the policies to overcome the global warming. The Earth Simulator has been built, as one of the policies, in order to obtain fundamental simulation data to predict global weather 100 years from now. Figure 3 is a snap shot of the global warming simulation carried out on the Earth Simulator [3].

The accuracy of weather simulations has been surprisingly improved by using fine meshes of 10 km. This was made possible for the first time by the outstanding performance of the Earth Simulator. Refinement of the weather simulation models is ongoing in order to obtain more reliable results about global warming. As can be seen from this example, simulations are becoming more influential than ever, not just in science and technology but also in human society.

4. Development of Next Generation Supercomputer at Greater-Than-Peta Scale
As mentioned above, following the decision by the Japanese Government to develop a next generation supercomputer, RIKEN was nominated to be the body responsible for conducting the project with funding from MEXT, since RIKEN had recent experience of developing an advanced supercomputer system, called the RIKEN Super Combined Cluster (RSCC), and is very capable of doing many kinds of cutting-edge scientific research.

To stimulate the active use of the supercomputer, a new law was enacted for the framework of its development and operation last year. Under this law, this facility will be open to all researchers from universities, research institutes and industries at home and abroad for both fundamental and industrial research, after the supercomputer is in full operation at the end of FY2011.

Figure 4 illustrates the current organizations involved in the project. As seen from the figure, the basic policies on the development of this project are handled by OSDP of MEXT. RIKEN is conducting the project as the body which develops the hardware architecture of the supercomputer, file systems, operating systems and so on in collaboration with Japanese hardware companies, national institutes, and universities. RIKEN is also developing the application software in collaboration with universities and other institutes.
The objective of the project is “Development, installation and application of advanced high-performance supercomputer system, as one of Japan’s ‘Key Technologies of National Importance’ (National Infrastructure)”. The following results are expected:

1) An advanced high-performance supercomputer system with the capability to carry out greater-than-peta scale simulations will be newly developed and be in full operation by the end of FY2011.

2) Application computer codes for utilizing the next generation supercomputer fully and efficiently will be developed and widely used.

3) Flexible computing environment will be provided by sharing the next generation supercomputer through connection with other supercomputers located at universities and research institutes.

4) An Advanced Computational Science and Technology Center (tentative name) will be founded as a base of computational science in Japan.

To achieve these four objectives, the following categories are considered in this project, that is, development of the hardware system, NAREGI as a middle ware [2], application programs of nano and life sciences as grand challenge. The image of idea is shown in Figure 5. It is important to recognize that this project has been planned not only for research activities but also from the industrial point of view.

The total budget for the project is roughly one billion US dollars, which includes the cost of the development of the supercomputer and, application software. It is now in the second year of its six year term. The status of the project is presently shifting from the basic design of computer systems to the detailed design of the computer chips. Several application programs are also being developed as grand challenges, which will be used to demonstrate the hardware performance and to stimulate computational science in Japan.
5. Basic Design of Computer System

First of all, the Committee of Application Software Investigation [4] was established in NSC, with more than twenty experts on computational science from universities, national institutes and industry in Japan in order to explore the best hardware architecture from the simulation point of view. Then, twenty-one application codes out of more than a hundred ones proposed from universities and research institutes all over Japan were chosen, since they have core algorithms to be used in the future supercomputers and are appropriate to investigate the basic hardware design. The research areas from which these codes were taken are nano-science (6), life sciences (6), environment and disaster protection (3), engineering (4) and physics and astronomy (2). The numbers in parentheses are the numbers of codes chosen. The computer companies and national institutes carried out desk benchmarks using these codes in order to search for the basic hardware concepts.

As a result, the two architectures were proposed, one by Fujitsu and the other by the joint group of Hitachi and NEC. A careful evaluation of these proposals is now underway by both RIKEN and the Japanese Government. Furthermore, the proposal made from RIKEN to MEXT has been authorized at the evaluation committee set up by MEXT as: 1) the hardware system is to consist of scalar and vector processor units, 2) the target performance with Linpack is to be 10 peta flops and 3) Fujitsu, Hitachi and NEC are to join the project as hardware developers [5]. Figure 6 gives an image of the next generation supercomputer proposed in this project. This proposal will be evaluated again at CSTP for the final review and be authorized to go on to detailed design of the computer chips.

As is widely known, Fujitsu produces supercomputers based on scalar architectures, and NEC has its own original technology for vector computers, which is employed in the Earth Simulator. However, as can be found in their proposals, it should be emphasized that many new advanced technologies have been invented to match the performance requirements for electricity consumption, floor size, etc. In other words, this project has already influenced hardware architecture and will lead future trend in supercomputer technology. This will become more evident as the project proceeds.

Figure 5. Schematic illustration of expansion of the developed hardware technology to information society
6. Major Applications as Grand Challenges

As mentioned above, the third S&T Basic Plan has pointed out the importance of the four research fields, that is, life sciences, information and telecommunications, environmental sciences and nanotechnology/materials. Following this policy, two of them, i.e., life sciences and nanotechnology/materials have been taken up to be the major target fields of the next generation supercomputer. Since the system will be highly parallel, computer codes for these simulations must be well adjusted to the new highly parallel architectures, and therefore careful tuning of the codes is greatly required, by synchronizing the development of the hardware itself.

Institute of Molecular Science (IMS), and RIKEN have been selected as the bodies to carry out these two application software development projects, IMS for nano-sciences and RIKEN for life sciences, and they are already in action as strategic footholds.

6.1. Nano-science

For the nano-science project by IMS [6], three main objectives are proposed as follows:

1. Simulations of Chemical Reactions of Enzymes

Simulations of the chemical reactions of enzymes are helpful in designing new artificial enzymes which have better reaction rates than natural ones. Here, theoretical frameworks based on quantum mechanics, along with 3D-RISM, are employed to take solvation effects into account, which is essential to describe chemical behaviors of the enzymes in water. The goal of this research is to design artificial enzymes which can produce hydrogen from timbers taken from demolished houses. Hydrogen is currently produced from biomass using corn, sugarcane and so on. Since these are foodstuffs, their use as raw materials has already affected food markets, therefore, used timbers are better raw materials in many ways. One difficulty in enzyme simulations is attributed to the environment in which their chemical reactions occur. Since reactions generally occur in water, free energy must be calculated to estimate correctly the activation energies of the reactions. 3D-RISM is considered to be a sophisticated theoretical scheme for free energy evaluations of the solvents. This approach will be helpful in searching for innovated technologies to overcome the energy issues recently well realized.

2. Simulations of Bio-Related Materials

In this research, theories and effective computational algorithms are to be developed for simulations of bio-related compounds such as proteins, cell membranes and viruses. Since the efficiency of these materials is extraordinary in biological environments, understanding of their mechanisms at a molecular level is one goal of this project. For example, ion channels in cell membranes let only specific ions like calcium and potassium pass through. The mechanism is not
well understood. The thermal motions of virus capsid are expected to reveal the detailed mechanism of virus infections to human cells. The methods to be employed include molecular dynamics, Fragment Molecular Orbital Theory, Monte Carlo scheme, and 3D-RISM. Consequently, so-called multi scale simulation is key and is considered to be a legitimate approach for biomaterial simulations. This approach is also taken in Subject 1, since enzymes function in solution.

3. First Principle Simulations of Nano-materials

The basic theory for this subject is Density Functional Theory in order to predict, for instance, the effects of impurities on semiconductors, which will be helpful to design new electronic devices. Models for solids to be treated in simulations contain more than one hundred thousand atoms to reproduce real electronic devices inside the computer. It is said that it takes about two weeks for researchers to examine the effect of impurities on devices by experiment, but only a few hours to probe such effects on a next generation supercomputer. Therefore, simulation is believed to improve the cost effectiveness for device development. Apart from computer devices, non linear optical materials and magnetic materials are also targeted in this subject.

To accomplish these objectives, IMS has been organizing research groups consisting of researchers from universities and institutes all over Japan. Furthermore, IMS has a close collaboration with RIKEN for careful tuning of the codes developed at IMS. This attempt is unique and is expected to produce highly efficient parallel codes for the next generation supercomputer. Figure 7 shows the basic concept for simulations in Nano-Science.

6.2. Life sciences

On the other hand, in the development of application software in the life sciences which is being organized by RIKEN, the four following subjects are being targeted to obtain knowledge for human health:

1. Molecular Simulations on Complexes of Protein and Ligands

   Behaviors of complexes of proteins and ligand molecules are theoretically studied here to understand the unique and efficient mechanisms of chemical phenomena occurring in a biological
environment. To evaluate free energies for these complexes in water, high accuracy ensemble calculations are needed. Here, hybridization of quantum mechanics, molecular mechanics and reaction field theories is key to establishing theoretical framework to simulate these complexes. By using effective ensemble methods based on these theories, the thermal motions of proteins, protein-protein interactions and protein folding will be calculated. The knowledge obtained from these simulations will be useful in designing new medicines. This approach is called in-silico drug design, which, it is believed, will become a standard technique to develop new drugs in their early stage of production. Protein-protein interaction is an interesting subject, as extremely large scale simulations which are only possible on a next generation supercomputer are needed.

2. Cell Simulations for Systems Biology
Systems biology is a new research field which was recently proposed to understand biological phenomena as a system. Motions of organelles are described considering non-uniformly distributed materials in each organelle and also changes of cell forms. Cell simulations are considered to be a core for systems biology. This will be a real challenge for the next generation supercomputer, since the theoretical models for cell simulations have not yet been clearly established. To overcome this difficulty, close collaborations with experiments need to be organized in this project.

3. Simulations of Human Organs and Whole Body
The basic mission of this group is to provide useful information for medical treatments. The detailed structure of the human body has been constructed inside computers using CT (computed tomography) technologies. Figure 8 shows a picture of a human body obtained by CT at RIKEN.

Figure 8. Human body structure measured by CT (DRI Computational Biomechanics Unit, RIKEN)

Based on this precise data, simulation of blood flow will be carried out to study how thrombus is created in the blood vessels. Recently, the use of supersonic waves to cure cancer is attracting attention, since it minimizes the damage to the human body, and especially in the case of brain cancers its effectiveness does not need to be explained. But, its problem is how to achieve pinpoint focus of the supersonic waves onto the tumor cells. Here, it is planned to analyze transmission patterns in the human body with simulations to search for focus points. Simulations of heart beats and air flow in the lungs are also involved in this group.

4. Bio-informatics to Relate DNA and Biological Functions
The study of genome sequences which are the origin of elaborate biological functions is crucial for life sciences. A major goal of this group is to establish fundamental research bases for tailor-made medical treatments using bio-informatics. For example, the effect of drugs differs from person to person because of single nucleotide polymorphism (SNP). Thus, if the relation between SNP and drugs becomes clear, the most efficient and safest dosage will be achieved.

**Figure 9 shows the basic concept for simulations in Life Sciences.** One of the unique aspects of the next generation supercomputer project is that it involves application software development from the beginning, which has not been common in Japan, but which is vitally important to achieve the targeted performance of the next generation supercomputer, since it will be a highly parallel computer.

![Figure 9. The basic concept for simulations in Life Sciences](image)

7. Schedules and Location

**Figure 10 shows the brief roadmap of the project.**

<table>
<thead>
<tr>
<th>FY</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGD</td>
<td>Operation</td>
<td>Start Operation</td>
<td>Final Operation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Software</td>
<td>Grid/Cloud middleware Design &amp; Production</td>
<td>Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Next Generation Nano-Science Simulation, Design &amp; Production</td>
<td>Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Challenge Application Software</td>
<td>Next Generation Life Science Simulation, Design &amp; Production</td>
<td>Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardware</td>
<td>Basic Design</td>
<td>Detail Design</td>
<td>Production</td>
<td>Enhancement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>File Systems and others</td>
<td>Design</td>
<td>Production</td>
<td>Enhancement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geographical Investigation, Construction</td>
<td>Investigation</td>
<td>Design</td>
<td>Construction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 10. Roadmap of the next generation supercomputer development**
The basic design of the supercomputer system has almost been finished and the detailed design of the computer chips will soon begin. From the next fiscal year, implementation of the chips will begin. Hardware production and installation is planned to be completed by the end of FY2011, including an enhancement to realize the targeted performance. In cooperation with the hardware design and production, application codes are developed both in nano-science and life sciences, which are supposed to achieve full performance on the supercomputer.

Last March, the city of Kobe, located about 600 kilometers west of Tokyo, was selected to be the location where the supercomputer center will be established. The supercomputer center will be the core of the Advanced Computational Science and Technology Center. **Figure 11 is an aerial photograph of the location.** After the establishment of the center, it is expected that collaborations with local governments and universities will prompt activities of educations, basic research, R&D use in industries and so on, using 40 giga bits network called Super SINET connecting between Tokyo and Osaka areas.

---

**Figure 11.** Aerial photograph of the location of the next generation supercomputer center

8. **Promotion Program of Supercomputer for Industries**

Almost at the same time as the project started, the Industrial Forum for Super Computing Promotion was established [7] with the participation of more than 160 Japanese companies to investigate how to use the supercomputer for industrial research and development. Its objectives are to use the next generation supercomputer for the following purposes:

1) Simulation of engines containing aerodynamics, combustion, structure and aging of materials and so on, which is named the Virtual Engine.

2) Analyses of car bodies aimed at production of cars without car prototype or crash and vibration tests.

3) Material simulation for the design of new functional materials, focusing on next generation LSI and nano technologies.
4) Polymer simulation for the design of new functional polymers for the future, although the design of polymers has become realistic to some extent now.

5) Weather simulation packages which have been widely used in applications for disaster measures caused by, for instance, storms and local downpours.

The Industrial Forum for Super Computing Promotion holds workshops every three months to inform industrial researchers about the usefulness of supercomputing. Delegates from the forum participate in the steering committee for the application software development organized by RIKEN, which makes it possible to have shared knowledge between academia and industry about next generation supercomputing.

9. Conclusions

The next generation supercomputer project has been launched to further stimulate computational science in Japan in order to motivate innovations of science and technology in both academia and industry. This is a Key Technology of National Importance endorsed by the third S&T Basic Plan and is being supervised by MEXT. RIKEN has been selected to be the body responsible for implementing the actual project under funding from MEXT.

The basic design of computer chips and the system is almost finished, and the detailed design will soon start. A unique aspect of the project is that nano and life sciences have been selected as important application areas for grand challenges and the computer codes are also to be developed in the project under collaboration with hardware researchers.

The supercomputer will be in full operation by the end of FY2011. Further, this project is expected to continue, even after the next generation supercomputer is in normal service, because successive development of supercomputers will keep Japan in a world leading position in science and technology both in computer device and simulation technology.

The science of the 21st century must tackle difficult and complicated problems for human survival and for the future of the earth. An integration of science that transcends the boundaries of different fields, the development of new sciences and innovation are needed. Therefore, in order to promote such new sciences, supercomputers capable of over peta scale computing are indispensable, and collaborations among nations concerned in development of computational science are highly desired.

Acknowledgments

The author is grateful to the staffs of Office for Supercomputer Development Planning of MEXT, and also the staff members of Next Generation Supercomputer R&D Center and Research Program for Computational Science of RIKEN and Institute for Molecular Science and JAMSTEC for providing the useful materials shown in this paper and also helpful discussions for the presentation at the SciDAC2007 conference.

References