4

Trends in Grid Technology
— Will the Grid technology become the core technology for the next-generation Internet application? —

MASAO WATARI
Information and Communications Research Unit

4.1 Introduction

Recently, the term "the Grid" has been seen more and more often. Grid technology makes computing resources available as needed, anytime anywhere through networks. They are called "the Grid" by the analogy with the electric power grid on which people can freely use electricity from wall outlets without being aware of generators.

There have been a number of examples that drew attention for their application of Grid technology to achieve outstanding performances. One example is that it enables to get one enormous computing power by using a network of high-performance computers, and another is that it enables to get supercomputer power by gathering the idle time of many PCs via the Internet. The latest research focuses on technologies to share not only computing power but also data storage and large-scale specialized laboratory equipment, which attract attention as an essential research infrastructure especially in "Big Science," such as high-energy physics and space science. As the Grid is also spotlighted as an effective tool in cross-disciplinary fields generated by the convergence of information technology and biotechnology / nanotechnology, many Grid application projects have started in the U.S. and Europe. It is said that the Grid will change the way of scientific research advancement.

Considering the Grid technology enables to realize one-step higher Internet applications, researchers, engineers, computer manufacturers and software vendors throughout the world are vigorously pushing ahead with researches, standardization and commercialization of Grid technology.

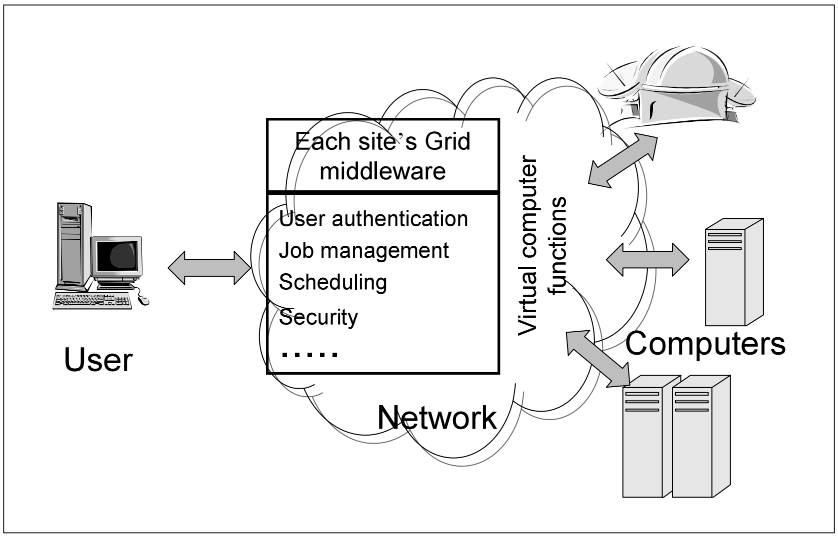
This article outlines noteworthy Grid technology and projects while describing Japan's strengths and weaknesses in this field, followed by the challenges Japan should address.

"The Grid," a term coined in supercomputer application technologies, was known as Grid Computing in the early days. As the definition of the Grid expands beyond mere computing, the word has become more inclusive, even referring to application environments in which large data bases and specialized experimental equipment can be shared. Applying a broader definition, these schemes are collectively referred to as "the Grid" in this article.

4.2 Vision of Grid technology

Like the utility model for power supply, the Grid is a concept that intends to provide environments in which computing and information resources are made available to users whenever necessary via networks. More stringently, the Grid is defined as "an environment where diverse computing and information resources (computers, storages, visualization devices, and large-scale experimental equipment) distributed across a network can be used as a single virtual computer by members of a virtual organization." A conceptual image is shown in Figure 1: computing and information resources can be used via network as a virtual computer for users registered as members of a virtual group formed across real organizations for a certain objective. The virtual computer's functions are realized through Grid middleware[*1] embedded within computers at each site. This allows users not only to perform any desired computations by

Figure 1: Conceptual image of the Grid



using their own program (computing services), but also to obtain the necessary processed data by using superior programs and data available from other sites (application services).

There are many benefits in creating a Grid. First of all, it provides scientists and engineers with a tool that enables them to efficiently conduct collaborative works from dispersed locations. In the Data Grid, a project funded by the EU, for example, 3,000 high-energy physicists jointly develop data and programs that are shared among them to encourage cooperation and competition of their research activities in order to improve their research efficiency. Since big-science research projects increasingly require massive-scale experimental equipment and data bases and highly intensive computations, sharing research resources is demanded to enhance the efficiency of research activities.

Secondly, Grid technology provides effective use of distributed resources with improved ease of use. For example, the computing power equivalent to a supercomputer can be obtained by gathering idle resources distributed in a network. Very large capability that would not be attainable by a single organization can be shared by connecting distributed resources via network. To take advantage of the high-speed computing capability, however, a high-speed network to ensure fast data transfer is indispensable.

Thirdly, additional benefits are load sharing and improved reliability. Intensive processing loads can be shared over distributed resources to avoid concentrated load on a single computer. This frees each computer from the need to prepare for peak load volume for the system because the overall networked resources complement the necessary capacity. Moreover, a failure in a single computer would not suspend the operation of the entire system, ensuring higher reliability. Even if a certain part of the system is struck by a cyber attack, there is no discontinuity in services, which are supported by the remaining system only with disconnection of the affected part, providing an effective solution to risk management.

Note that the Grid is not a technology to advance a computer itself, but is instead a computer application technology. There are misunderstandings such as the Grid can substitute as a large supercomputer or a large supercomputer can substitute as the Grid. Efforts to raise the performance of supercomputer itself are still needed because there are many computation applications that are significantly inefficient if performed over a network. Improving the performance of individual high-performance computers benefits the Grid gaining larger computation capability.

4.3 History of the research on Grid technology

Grid technology research derives from the studies on remote computing and distributed computing in the 1980's. Driven by the high-speed networks that became available in the 1990's, not only theoretical studies but also building their demonstration systems appeared.

In 1992, Charlie Catlett and Larry Smarr, the National Center for Supercomputing Applications (NCSA), the U.S. published a concept called "metacomputing."[1] They started research in order to realize a virtual computer environment on a network for large-scale parallel computation. In 1995, Tom De Fanti, the University of Illinois, Rick Stevens, the Argonne National Laboratory and others performed the first large-scale metacomputing experiment under the I-Way (Information Wide Area Year) project. In this project, 17 computer centers nationwide were linked together via a high-speed wide-area network for the demonstrations where many applications such as virtual-reality experiments were realized. Originating the I-Way, the Globus Project[2] launched by Ian Foster, Argonne National Laboratory and Carl Kesselman, the University of Southern California in 1996. They developed the middleware for high-performance distributed computing, followed by the 1998's publication of the blueprint describing the concept of "Grid technology."[3] The Gird was defined as "an environment in which users can use computing resources as needed at anytime without being aware of the location of the computers."

The Globus Toolkit, the software developed in the Globus Project, offers basic Grid middleware capabilities such as user authentication and Grid resources management. The software, which is made available by Argonne National Laboratory as open source, has been utilized in so many research projects for Grid constructions that it has become the de facto standard. Building a Grid system, however, requires the development of a number of programs in addition to the Globus Toolkit, which provides only the basic functions.

There is another direction of research aiming to make good use of idle CPUs through the Internet. In 1985, Miron Livny, the University of Wisconsin proposed an idea of utilizing distributed workstations' idle CPUs for computations. In 1991, he succeeded in gathering the equivalent of 400 CPUs in a project known as Condor. In 1997, Scott Kurowski established Entropia, Inc., and in two years latter successfully collecting the computing power of idle PCs through the Internet could provide Tera*2 flops level computation, which is almost equivalent to the highest level of performance delivered by a supercomputer of that time.

Metacomputing and the utilization of idle processing power, both of which provide virtual single computer environment through networks by distributed computers, are component technologies of the Grid.

In Japan, meanwhile, research on global computing, a technology to allow people to operate computers from remote places, is in progress. One of the activities toward this end is Ninf (Network Infrastructure for Global Computing), a project launched in 1994 by Satoshi Sekiguchi, the Electrotechnical Laboratory (currently known as the National Institute of Advanced Industrial Science and Technology), Satoshi Matsuoka, the Tokyo Institute of Technology and others. As global computing research, they worked on the design and experimental proof of Remote Procedure Call (RPC), protocols between clients and servers enabling that clients make servers in the network perform computations. The project's outcomes are now under review for standardization as part of Grid middleware. Researchers involved in Ninf are now collaborating with the participants of NetSolve, a similar independent initiative of the University of Tennessee, to make further advances in research and development. In addition, the Japan Atomic Energy Research Institute, the Institute of Physical and Chemical Research, and computer centers at universities have contributed to the studies on technologies to enable the remote or shared use of supercomputers via networks. These activities have led to the creation of the essential basis for the development of Grid technology.

4.4 The background of the emergence of the Grid

Major driving forces for the recent considerable development of the Grid are the growth of the seeds of technology and the rise of needs of applications. In technology developments, the latest Internet backbone networks provide 10-Gbps level for domestic access and 1-Gbps level for international access, ushering in the so-called broadband age. Furthermore, network speeds and reliability that have been improved along with the growing the Internet allows the Grid to provide easy-to-use computing environment for distributed computers via the Internet.

In the domain of application needs, the Grid is beginning to be viewed as an indispensable research tool for IT-driven activities in science, such as biotechnology and nanotechnology represented by e-Science.[*3] On the other hand, big science such as high-energy physics and space science intends to improve research efficiency by sharing expensive specialized laboratory equipment and data analyses, positioning the Grid as an infrastructure for research activities.

4.5 Applications of the Grid

Among the variety of possible applications of the Grid, typical examples expected in the future are described below.

(A) Metacomputing

Performing large-scale computations that are impossible on a single supercomputer by simultaneously running multiple high-performance computers (HPC) such as supercomputers distributed across a network. The intent is to create a virtual enormous computer. There are two possible types of computations performed on the Grid: single-program computations that are divided and distributed to parallel computers; and computations for which the same program is installed onto multiple computers so that each computer is given a different set of data in parallel processing with collecting the results at each computer (parameter sweep). Since data transfer rates on networks are slower than those inside a computer on the order of magnitude, computations performed on the Grid should be of a small data access range, or so-called small grain computations, to make the most of its capabilities.

(B) Research Grids (Virtual Laboratories)

Building a community of researchers and research institutions, each participant can access community's resources of computers, databases, and experimental equipment through networks. In addition to using each other's data, this scheme allows researchers to couple their application programs with those of others to perform complex simulations.

In conventional remote access, each site had its own software. The Grid provides common interfaces for remote access enabling broader connectivity and easy-to-build.

Research Grid activities in the U.S. include, as shown in Table 1-1, Grid Physics Network in high-energy physics, Fusion Grid in nuclear fusion, the National Virtual Observatory for astronomical observation, the Earth System Grid in meteorology, and the Biomedical Informatics Research Network (BIRN) in bioinformatics, which are also seen in Europe through similar projects.

(C) Access Grids

Access Grids provide an environment that facilitates smooth and speedy operation of joint research projects by enabling remote researchers to share the same screen views and computational results. In addition to an image quality better than TV conference systems through multicasting via high-speed networks, Access Grids offer tools such as file sharing for efficiently conducting collaborative work.

(D) Data Grids

Also known as Data Intensive Computing, Data Grids allow network-based remote access to data that is too large to be stored in a single location or data that is distributed to physically separate locations. The technology is still under study with such themes as efficient storage and readout of large data sets, and high-capacity Internet communications. The key here is the data transfer time. Since the transmission of large volumes of data takes a considerable time even over a super-high-speed network, primary analyses and calculations need to be performed on the site of data generation.

In an experimental project on a huge accelerator at the CERN (European Organization for Nuclear Research), for example, 7,000 researchers from 500 institutions perform several hundred million times of experiments annually with producing 6-8 petabytes of data. In order to analyze such experimental data on high-energy physics and to share the results among researchers, the Grids are under development in projects such as the EU Data Grid in Europe and the Grid Physics Network (GriPhyN) in the U.S. Grid technology is also used for the space research to describe the entire universe based on the data obtained from more than one space observatory. Thus, Data Grids improve the efficiency of activities in big science.

(E) Computer Service Grids

Providing as much computing power as needed on an on-demand basis without the users being aware of the type of computation server, by a Grid having multiple computation servers on the network within an organization. A Computer

Table 1-1: Major Grid projects in the U.S.

Project	Participants	Budget	Term	Objective
PACI National Technology Grid	NCSA, SDSC	NSF	1998~	Dissemination of Globus and creation of a computational grid to support dispersed laboratories
Information Power Grid	NASA	NASA ($7-8M for application, excluding infrastructure)	1999~	Implementation and visualization of multi-disciplinary simulations on a heterogeneous distributed computing system
Access Grid	ANL, LBNL, LANL, NCAR, NCSA, etc.	DOE,NSF $2M/Y	1999~	Development and deployment of Grid-based remote conferencing systems and other collaboration systems. Part of the outcome was used to broadcast GGF meetings to Access Grid members.
ASCI Grid (DISCOM)	Sandia NL, LLNL, LANL	DOE	1999~	Application of ASCI to the Grid, Security development
Grid Physics Network (GriPhyN)	ANL, U of Florida, Fermi Lab, etc.	NSF $12M/5Y	2000~	Sharing data in high-energy physics and astronomy
Particle Physics Data Grid(PPDG)	ANL, LBNL, Caltech, SDSC, etc.	DOE $9M/3Y	2000~	Sharing data in particle physics
Tera Grid	NCSA, SDSC, ANL, Caltech	NSF $12M@2001, $53M@2002	2001~	Construction of "One Peta byte Gateway": 13.6TF, 6.8TB of memory, 79TB built-in hard drive, 576TB network-aggregated hard drive
International Virtual-Data Grid Lab (iVDGL)	U of Florida, U of Chicago, Indiana U, Caltech, Johns Hopkins U, etc.	NSF $13.7M/5Y	2001~	Linking of US, European and Japanese Data Grids for technical development and expansion of the Data Grid
Network for Earthquake Eng. Simulation Grid(NEES)	U of Southern California, U of Michigan, ANL, NCSA, etc.	NSF $10M/3Y	2001~	Earthquake simulations
National Virtual Observatory (NVO)	Johns Hopkins U, Caltech,etc.	NSF $10M		Virtual space observatory
Grid (Grid Research Integration Deployment and Support) Center	U of Chicago, U of Southern California, U of Illinois, U of Wisconsin, etc.	NSF $12M	2001~	Integration of Grid software and deployment and support for Grid platform software
DOE Science Grid	DOE			The generic name for DOE's scientific Grids; Dissemination and deployment of the Grid; Development of support tools for researchers
Fusion Grid	ANL, LBNL, Princeton Plasma Physics Lab, General Atomics, MIT, etc.	DOE $6M/3Y	2001~	Grid for nuclear fusion research
Earth Systems Grid	ANL, LLNL, NCAR, etc.	DOE $5M/3Y		Grid for meteorological research
Biomedical Informatics Research Network		NIH $20M	2002~ 2006	To develop technologies for visualization of the brain and the neural system and to share databases

ANL=Argonne National Lab, LANL=Los Alamos National Lab, LBNL=Lawrence Berkeley National Lab, LLNL= Lawrence Livermore National Lab, NCSA=National Center for Supercomputing Applications, SDSC=San Diego Supercomputer Center, NCAR=National Center for Atomospheric Research

Source: Author's compilation from relevant web sites.

Service Grid works like a virtual computer center. An example is the Campus Grid of the Tokyo Institute of Technology. It consists of PC clusters (a total of approximately 800 processors) distributed campus-wide and 25 terabytes of storage that are both accessible via the high-speed campus network. Another example is seen in the private sector, where companies have constructed Computer Service Grids by connecting their computer centers through their Intranets.

While the commercialization of Computer Service Grids is under discussion as an extension of this technology, there remain problems such as pricing and security. Therefore, current applications are limited to closed networks built on campus or corporate Intranets.

(F) Grid ASPs (Application Service Providers)

Grid ASPs offer services in which remote users send data through a network to be run on application programs operating on a high-performance computer, and receive the results. While Computer Service Grids mentioned before require users to make programs by themselves, users of Grid ASPs can make use of ready-to-run programs with proven functionality. Another potential service is to provide useful databases such as genome data to Grid ASP users. With the Grid, medical data that may not be exposed for privacy protection, for example, could be made available on the limited condition that the data is processed within the site of the database provider by using a desired analytical program so that only the results are sent out to the Grid ASP users.

(G) Desktop Grid Computing

Being aware that home-use PCs sit idle most of the time, this scheme intends to gather the untapped computing power of these PCs in order to perform certain computations. It is also known as Volunteer Computing as participants usually offer their idle computing power for free.

The technology has been applied for space observation data analysis (SETI@home: Search for Extraterrestrial Intelligence at Home) and for the development of new drugs to cure cancer, AIDS and leukemia (Parabon's Compute-against-Cancer, Entropia's Fight AIDS At Home, United Devices' Cancer Research Project, and Intel's Philanthropic Peer-to-Peer Program). In the case of SETI@home, four million volunteers offer their PCs' idle time to achieve the computing power equivalent to a 40-TeraFlop supercomputer.

In Desktop Grid Computing, only limited types of computations are possible because the processing power of each participating computer is not high enough and these computers are linked via slow networks. Parameter sweep applications that execute the same process on large volumes of different input parameters, are ideal for this type of Grids. It is necessary for the Grid to have software upgrades, security, privacy protection and fault tolerance (ability to resume in case of failure).

Recently, some companies have built systems for order/purchase management by using the idle computing power of their PCs. In their cases, as the PCs are connected to their own Intranets, there is no need for concern about security and fees.

(H) Sensor Grids

A Sensor Grid is one of the ultimate form of the Grid in ubiquitous computing in the future. Many Sensors installed everywhere are connected to the Internet to allow access to data from any place. A possible application is the monitoring of the global environment by collecting and analyzing data obtained from numerous sensors positioned all over the world including remote places.

4.6. Grid projects

Based on the awareness that the Grid not only serves as a critical infrastructure for the advances of big science, but also provides an indispensable basis for the convergence of IT and biotechnology / nanotechnology, the U.S. and European countries are actively conducting national projects on the Grid. The following sections describe the activities in the U.S., Europe and Asia, which are listed in Table 1-1 to 1-5.

(A) U.S. projects

Since Grid technology originates from a study in the U.S. at the use of supercomputers via a network, the U.S. already has an ample accumulation of resources from the years of

Table 1-2: Major Grid projects in Europe

Poject	Participants	Budget	Term	Objective
Grid in 6th Framework Program		EC 300 M Euro/5Y	2003-2007	FP6 covers IT, biotech, nanotech, astronomy, food, environment, energy, and intelligent society. Grid construction is planned in each area with a total investment of 300M Euro.
EU Data Grid	CERN, U of Heidelberg, IBM UK, CNRS, INFN, PPARC, SARA, etc.	EC 9.8 M Euro/3Y	2001-2003	Development for realtime execution of petabyte data processing via Grid network
EuroGrid	Forschungszentrum, Pallas GmbH, U of Bergen, CNRS, Warsaw U, U of Manchester, ETH Zurich, etc.	EC	2001-2003	Grid applications development: molecular biological modeling, meteorological forecasting, CAE simulations, and Grid middleware development
GRIP (Grid Interoperability Project)	U of Southampton, Deutcher Wetterdienst, U of Manchester,, Pallas GmbH, U of Warsaw U, etc.	EC	2002-2003	Building up compatibility between UNICORE and GLOBUS
MAMMOGRID	U of West England, U of Pisa, U of Oxford, U of Cambridge, Mirada Solutions, etc.	EC	2002-2005	Development of Mammogram database for healthcare research

Table 1-3: Major Grid projects in the U.K.

Project	Participants	Budget	Term	Objective
eScience		£120M/3Y	2001-2004	The umbrella program to promote scientific research with Grid development. £75M of the budget is for Grid applications development.
Grid Particle Physics	Universities of Birmingham, Bristol, Cambridge, Edingburgh, Glasgow, Lancaster, Liverpool, Manchester, Oxford,Sheffield, Sussex, Imperial College, CERN, etc.	DTI PPARC £17M/3Y		Particle physics research jointly conducted by EU DataGrid(CERN), US GriPhyN and PPGrid
Astro Grid	Universitiies of Edinburgh, Leicester, Cambridge, Queens Belfast, etc.	DTI PPARC £5M	2001-2004	Astronomical observation and research jointly conducted by EU AVO and US NVO.
DAME (Distributed Aircraft maintenance Environment)	Universitiies of York, Oxford, Sheffiled, Leeds, etc.	DTI EPSRC £3M	2002-2004	Sharing of aeronautical data and engine data for aircraft maintenance
Reality Grid	Universitiies of London, Manchester, Edinburgh, Loughborough, Oxford, etc.	DTI EPSRC £3M		Tools for materials research
myGrid	Universitiies of Manchester, Southampton, Nottingham, Newcastle, Sheffiled, etc.	DTI EPSRC £3M	2001-	Creation of environments in which biological and other data are made accessible to medical doctors as well as to IT engineers
Biology, Mediacal, Environmental Sciene	Under consideration	DTI £23M		Grid in biology, medical and environmental science fields

DTI=Dep. of Trade and Industry, PPARC=Particle Physics and Astronomy Research Council, EPSRC=Engineering and Physical Sciences Research Council, NERC=Natural Environment Research Council

Source: Author's compilation from relevant web sites.

Table 1-4: Major Grid projects in Japan

Project	Participants	Budget	Term	Objective
IT Based Laboratory (ITBL)	Japan Atomic Energy Research Institute, Institute of Physical and Chemical Research Japan, etc.	MEXT ¥5B ¥5.2B@2000 ¥2.7B@2001	2001-	Creation of a virtual research laboratory environment by linking supercomputers via Super SINET
Research funded by Grant-in-Aid for Scientific Research on Priority Areas “Informatics” A05Grid	Osaka U., Tokyo Institute of Technology, National Astronomical Observatory Japan, etc.	MEXT ¥0.8B /5Y	2001-2005	Medical Grid by Osaka U.’s Shimojo Group, P2P Data Grid by TITech’s Matsuoka Group, Nationa Astronomical Observatory Japan, etc.
TITech Campus Grid	Tokyo Institute of Technology	FY2001’s supplementary budget ¥0.2B	2001	Construction of a Grid on the campus
BioGrid Center, Osaka U.	Osaka U.	FY2001’s supplementary budget ¥0.2B	2001	Construction of a Grid infrastructure
e-Science project “Construction of a supercomputer network”	Osaka U., etc.	MEXT ¥0.55B @2002	2002-2006	Development of technology to link large-scale Data Grids and Computing Grids
Development of network computing technology “Grid Cluster Federation”	National Institute of Advanced Industrial Science and Technology, Tokyo Institute of Technology, U. of Tsukuba, U. of Tokyo	METI 1.3B/5Y	2002-2006	Realization of cluster technology, Grid technology and 10-petabyte-level massive data processing
National Research Grid Initiative (NAREGI)	Under consideration	MEXT (under consideration)	2003-2007	Creation of the world's highest level Grid environment to promote interdisciplinary research through a marriage of biotech / nanotech with IT
Grid computer “Business Grid Computing”	Under consideration	METI (under consideration)	2003-2005	Development of software that provides flexible services by using numerous networked computers/storage systems like a single computer

Source: Author's compilation from relevant web sites

Table 1-5: Major Grid projects in South Korea and China

Project	Participants	Budget	Term	Objective
Korea Grid Initiative	Led by the Korea Institute of Science and Technology Information (KISTI)	¥4.3B/Korean Ministry of Information & Communication (not including Network)	2002~2006	Development of Grid Middleware; Data Access; Applications development (over half of the budget is for applications development)
China Computational Grid	Chinese Academy of Science	$40M/Ministry of Science and Technology		Connection of 10 HPC centers via Network; One of applications is Genomic Aapplications.

Source: Author's compilation from relevant web sites

developing Grid technology and of Grid demonstration systems. The nation continues in its vigorous activities for further promoting the development of Grid technology, large-scale demonstration systems and Grid applications.

In the NSF’s TeraGrid project, as one example, research on processing large volumes of data on the order of terabytes is in progress. Four computing centers nationwide are constructing a large-scale integrated demonstration system in which high-performance computers with a combined processing capacity of 13.6 teraflops and an enormous storage space of 576 terabytes are planned to be connected to a 40-Gbps high-speed network. Another noteworthy initiative is NIH’s Biomedical Informatics Research Network (BIRN). Through this project, 10 medical research institutions nationwide, with the University of California, San Diego as the leader, are involved in the development of a Grid to visualize the activities of the brain and to share the acquired data, in order to study Alzheimer and other diseases. A large number of researchers, including medical science and medical care, participate in the program.

Meanwhile, since the September 11 terrorist attacks, the U.S. government has been discussing measures to tighten homeland security, of which the defense against cyber terrorism is a key element. From this point of view, security of their Grid projects is being reviewed, giving higher priority to security issues in the Grid world.

(B) U.K. and European projects

In the U.K., a program known as e-Science was launched in 2000. The program's intent is to encourage and accelerate scientific research activities by using IT. Specifically, the government is actively promoting Grid-enabled projects that pursue the sharing of expensive computers, laboratory equipment and large data sets distributed worldwide. As the nation's Grid infrastructure, they have established nine National e-Science Centers nationwide and they are building the Grid by connecting high-performance computers at via a high-speed network. Applications developed on this Grid range from high-energy physics, genomics, biotechnology, protein structure analysis , medicine and health to environment, meteorology, astronomy, chemistry and materials. Dr. John Taylor, Director General of the Research Councils OST UK, who leads the e-Science program, says that the Grid is a necessary infrastructure for the U.K. to participate in global scientific research activities.

Even without the presence of domestic supercomputer makers, European countries have the accumulated expertise in the application of supercomputers and have started activities toward Grid applications early. EU's representative project is the EU Data Grid, which is designed to handle large-scale data. The project, led by CERN, has been carried out in collaboration with the U.S. and Japan.

Since the EU-funded projects must be based on joint development of more than one EU member nation, Grid projects, which usually aim at the construction of a common research infrastructure, fit very much in the purpose of the EU project. Other ongoing European efforts for Grid projects except for those in the U.K. include Unicore in Germany, INFN Grid (Italian National Institute for Research in Nuclear Physics) in Italy, and Grid-based Virtual Laboratory Amsterdam (VLAM) in the Netherlands.

(C) Japanese projects

In IT-Based Laboratory (ITBL), a project started in 2001 that seeks to construct a virtual laboratory, the Japan Atomic Research Institute, the Institute of Physical and Chemical Research and others are involved in building a common research environment by supercomputers via a communication network. By fiscal 2001's supplementary budget fund, more Grid-oriented projects have been launched such as the establishment of Grid Research Center in National Institute of Advanced Industrial Science and Technology, and the construction of Tokyo Institute of Technology's Campus Grid. In fiscal 2002, the BioGrid Project led by Osaka University commenced, and the construction of the Grid for bioinformatics has started by the University of Tokyo, Japan Advanced Institute of Industrial Science and Technology, the University of Tokushima, the Institute of Physical and Chemical Research and others as part of the "Genomic Information Science" research project supported by the grant-in-aid for the scientific research. Furthermore, in fiscal 2003, full-scale Grid construction activities such as the National Research Grid Initiative, an effort toward the creation of the world's highest-level Grid, and the Business Grid Project, an attempt to develop software applications that enable flexible Grid services, are planned.

In addition, Super SINET, a project launched in 2000 and serving a high-speed network connecting computer centers at universities and national laboratories, is now an essential infrastructure for Grid construction

(D) Asian projects

In the Asian region, major countries that are conducting Grid research are South Korea, China, Taiwan and Singapore. Japan takes the initiative in regional Grid promotion through Asia-Pacific Grid (APGrid) and the Pacific-Rim Application & Grid-Middleware Assembly (PRAGMA), promotional organizations for Grid activities in the Asia-Pacific region including the U.S. and Australia. In addition, a high-speed network infrastructure for the Grid is being built through the activities of Asia-Pacific Advanced Network (APAN). These successful efforts have allowed Asia to secure a position as one of the world's three poles in the Global Grid Forum (GGF), a global standardization organization for the Grids. Asian nations are actively launching Grid projects to promote fast catch-up. Japan must make further progress to stay ahead of them.

(E) Challenges in Japanese projects

Projects in the U.S. and Europe place emphasis on the development of Grid applications as well as Grid technology. For applications development, cross-disciplinary collaboration between IT and biotechnology / nanotechnology is important. While it has been seemingly rare in Japan for researchers to actively participate in research in foreign fields, IT engineers are now asked to jump into Grid application areas to create software required in each of application areas.

There are many pieces of software born at universities in the U.S. and Europe. As this is also the case with the Grid, a number of Grid software ventures have already been spun out from universities. Based on the Legion project, launched at the University of Virginia in 1993, Prof. Andrew Grimshow established a company named Applied MetaComputing (now known as AVAKI) to commercialize the Legion's Grid technology.

On the other hand, in Japanese projects many pieces of software have been developed however there were very rare cases of commercialization based on result of projects by universities. The reason is that as projects at universities cannot afford many engineers, what they can do is limited to the development of a core program and the demonstration system which shows only essential component functions. Commercialization of their software requires a larger number of engineers to proof their functions by larger-scale experiments and to add peripheral programs for such as application interface.

In the projects at U.S. universities, which are enabled to employ more research engineers, larger-scale demonstrations are performed. The U.S. government often demands project members not only to conduct basic research but also to verify the practicability of the outcomes. Researchers who have been involved in such projects are given options such as establishing a venture business or participating in another project when the project ends. In Japan, where until fairly recently universities could not employ staff specifically for their projects, a lack of job mobility among research engineers and personnel shortfalls remain unchanged. How to handle intellectual property rights at universities is still in a transitional stage, although systems for this end are being established.

It is hoped that many pieces of software are created based on the results of large-scale projects at Japanese universities. In the U.S. and Europe, projects at universities are the birthplace of many innovative software applications, indicating that universities are the foundation of the nation's technical capabilities in the software field. In order to enhance Japan's technical strength in software development, expanding the capability of Japanese universities to create software systems and increasing their accumulation of technical expertise are essential.

4.7 Trends in standardization

Since worldwide connectivity is the key to taking advantage of the Grid, how to standardize its interface is very important issue in the Grid communities.

Activities for Grid standardization started with the formation of the Grid Forum (GF) consisting of researchers involved in Grid projects in late 1998. Subsequently, the European Grid Forum (E-Grid) and the Asia-Pacific Grid Forum (AP-Grid) were founded in Europe and Asia-Pacific respectively, with a view to promoting regional Grid development. In 2001, GF merged with the standardization groups in E-Grid and AP-Grid, resulting in the establishment of a global standardization organization known as the Global Grid Forum (GGF). With two Japanese members taking part in the Steering Group and one in the external advisory committee, Japan has shown its presence in the Grid community.

GGF adopts the same standardization process as the IETF (Internet Engineering Task Force), a

Table 2: List of groups organized for the Global Grid Forum

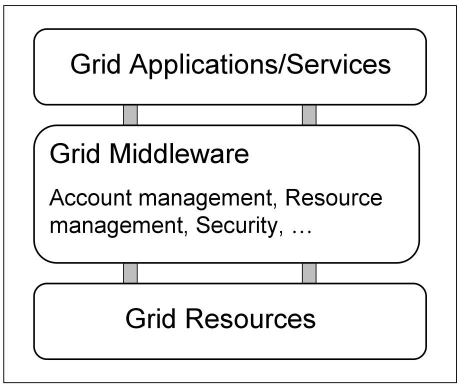
Area	Working Groups	Research Groups
Applications and Programming Environments		— Advanced Collaborative Environments (ACE-RG) — Advanced Programming Models (APM-RG) — Applications and Test Beds (APPS-RG) — Grid Computing Environments (GCE-RG) — Grid User Services (GUS-RG)
Architecture	— Open Grid Service Infrastructure (OGSI) — Open Source Software (OSS) — New Productivity Initiative (NPI)	— Grid Protocol Architecture (GPA) — Accounting Models (ACCT) — Service Management Frameworks (JINI)
Data	— GridFTP — Data Access and Integration Services (DAIS)	— Data Replication (REPL) — Persistent Archives (PA) — Grid High-Performance Networking
Information Systems and Performance	— Discovery and Monitoring Event Description (DAMED) — Network Measurement (NM) — Proposed: Grid Information Retrieval (GIR) — Proposed: CIM based Grid Schema	— Relational Grid Information Services (RGIS) — Grid Benchmarking (GB)
Peer-to-Peer	— NAT/Firewall — Taxonomy/Nomenclature — P2P Security — File Services — Instant Messaging Interoperability	
Scheduling and Resource Management	— Grid Resource Allocation Agreement Protocol (GRAAP) — Management Application API Working Group (DRMAA) — Scheduling Dictionary (DICT) — Scheduler Attributes (SA)	
Security	— Grid Security Infrastructure (GSI) — Grid Certificate Policy (GCP) — Open Grid Service Architecture Security (OGSA-SEC)	— Kerberos

Source: Web site of the Global Grid Forum at http://www.gridforum.org/L_WG/wg.htm

standardization group for the Internet. In other words, research groups and working groups in GGF are organized for each technology to discuss the relevant issues with clear objectives and goals. Each group will be dissolved when its mission is completed. At present, GGF has 21 working groups and 14 research groups in seven areas as shown in Table 2. Group members hold discussions through group meetings and exchange opinions via e-mail, in addition to the general meetings held three times a year, allowing speedy progress of the standardization discussions. For those who are abroad, keeping up with their activities requires considerable effort. The results of GGF have become the de facto standards in the field, but are not de jure or enforced rules.

The focus of Grid standardization activities so far has been Grid middleware as shown in Figure 2, which serves as common system software across application fields. Grid communities are competing to take the initiative in setting up an interface with higher-layer applications or with lower-layer Grid resources. As Grid technology covers a wide range of functions, dominance by a single software package or a new middleware package is unlikely. Therefore, Japan or any single company has no chance of attaining supremacy in standardization using its particular idea. Instead,

Figure 2: Layers of Grid system



Japan can contribute to the standardization by playing a complementary role in worldwide collaborative activities to advance superior technologies.

In June 2002, the Grid Consortium Japan was founded aiming at conducting the development and standardization of Grid technology under the coordination of standardization activities as well as the promotion of Grid technology and the exchange of its information. As the development from standardization to commercialization usually takes very short time in the field of Internet software, the activities in the standardization organizations must be carefully monitored.

4.8 Challenges to Grid technology

With Globus Toolkit and other types of middleware available, many projects today focus on the construction of Grids or the development of applications to run on the created Grids. However, the available software still falls short of enabling researchers to easily build Grid-based systems. Current Grid technology imposes a variety of restrictions that researchers intend to remove through their efforts.

Current challenging research themes include security control across different organizations, alignment of policies for operation, pricing schemes, actions at failure, ensured communication channels by QoS (Quality of Service), dynamic resource management, and large-scale data processing. More specifically, research subjects expected for the next three to five years are: security control between firewalls protecting from foreign access and Grid-enabled access from other organizations, large-scale data transfer in which a broadband communications channel must be secured through QoS control, dynamic resource management through the constant monitoring of online resources, resolution of the contradiction between computer operating policies that vary from organization to organization to provide an integrated operating policy for Grid users, and systems to charge fees for external access.

4.9 Recent trends among makers in Japan and elsewhere

Construction of the Grid requires high-performance computers and high-speed networks for the hardware infrastructure. In the U.S., active participants in the development of the Grid are IBM, Sun Microsystems and Hewlett-Packard (including former Compaq), with Microsoft giving attention to Grid middleware.

In March 2002, IBM proposed applying the standards for Web services to Grid middleware, an idea having been approved by GGF. This has quickly raised high hopes for advances in the commercialization of Grid technology, the scheme of the proposal suggests new approach for integration with e-commerce platforms. Thus IBM, in addition to the sales of high-performance computers for the hardware infrastructure, has been making considerable efforts for the dissemination of Grid middleware, the software products they are actively developing.

In Europe, where there is no computer manufacturer, hardware used for Grid construction consists of computers made by mainly U.S. makers. Accordingly, business activities toward the development of Grid applications are jointly conducted with U.S. counterparts.

When looking at the Asian region not including Japan, significant efforts to build Grids are being made there as well. With a recent leap in the performance of PC clusters and high-performance computation servers, the construction of a Grid is now relatively easy by deploying US-made computers.

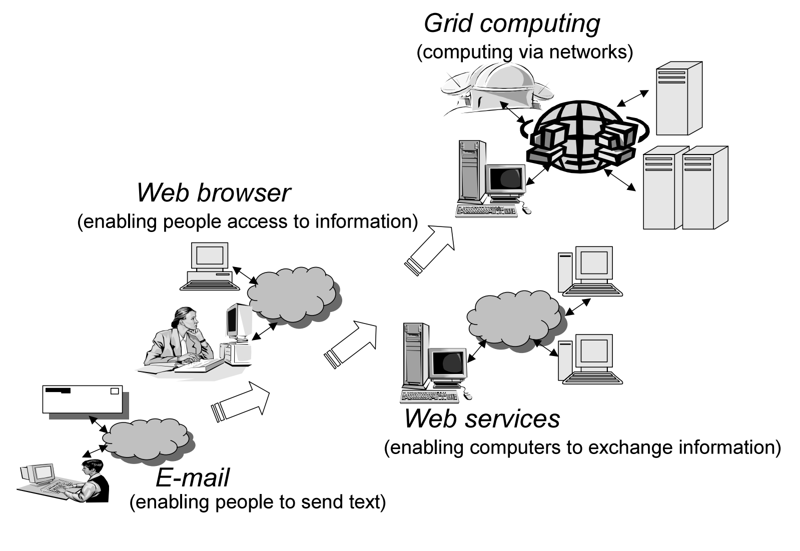
Meanwhile, Japanese computer makers, which are behind U.S. counterparts in the area of Grid development, are struggling to catch up.

4.10 Positioning Grid technology

Grid technology, as one of the Internet application technologies, appears to have the potential for significant growth when taking into account how Internet technology has evolved. The history of the development is shown in Figure 3.

Application of the Internet started with e-mail, a

Figure 3: Development of Internet applications



technology to exchange messages via networks. Once Web browsers appeared as a tool to view information on remote sites, their convenience led to the explosive expansion throughout the world. The growth was attributable to the advantage that Web browsers enable users to view information regardless of the types of terminal devices or operating systems they use. With its expanded availability and enhanced reliability, the Internet has become a common tool for a range of business operations.

Following the phase in which only humans could browse online information, Web services technology emerged and enabled computers to automatically retrieve and interchange information through a wider range of interaction with other computers. Web services have also changed the way of data exchange between computers, a technology known as EDI (Electric Data Interchange), by allowing data exchange over the Internet as an open network instead of the conventional methods based on platform-specific rules. The technology is expected to facilitate safe e-commerce and Business-to-Business transactions.

The Grid can be considered as an extension of Web services technology. While Web services enable access to data and information available on the Internet, the Grid allows the use of computing resources such as computers, databases and laboratory equipment connected to the Internet.

4.11 The future of the Grid and expected challenges

(1) Significance of the Grid

Grid technology enabling computer resources on a network to act a virtual single computer with improved ease of use, is expected to become the core technology in next-generation Internet applications. The Grid also provides a valuable research infrastructure for many fields by enabling the remote use of expensive computers, large-scale data, useful analytical programs, and high-cost experimental equipment. For example, in big science such as high-energy physics and space science, the sharing of high-cost computers, expensive laboratory equipment, and remote facilities can result in improved research efficiency. The Grid can also benefit cross-disciplinary research fields created by the convergence of information technology and biotechnology/nanotechnology, where data and analytical programs can be jointly developed and shared, promoting collaboration in research. The Grid will become an essential element of the research infrastructure in fundamental science, with the potential of changing the conventional styles of research.

The Grid thus can possibly encourage joint research projects, and a growth in joint activities in interdisciplinary areas may lead to the creation of new fields of study. Traditionally, the majority of

Japanese research activities were self-contained due to barriers that divided each research field and inhibited interdisciplinary cooperation between researchers from different domains. The promotion of such collaborative initiatives is another advantage of the Grid's construction.

(2) Toward the development of Grid technology

Born in the U.S., the concept of the Grid has developed mainly in the country of its origin. Europe, the region with no computer makers, has been deploying U.S.-made high-performance computers to actively develop Grid applications in collaboration with U.S. research institutions. On the other hand, fortunately, Japan has leading research expertise in Grid technology together with domestic computer manufacturers, an environment that allows the creation of prominent technologies through Grid projects jointly conducted by universities, national research laboratories and computer makers. Given that homeland security has been reinforced in the U.S. since the September 11 terrorist attacks, security must be considered as a critical factor in the development of Grid technology.

Standardization is another key issue in Grid technology development, as they act as Internet-enabled software. Grid technology is to be standardized through international collaboration aiming to create superior technologies, rather than through the dominance by a single company or nation with its particularity. In such a situation, Japan has to come up with outstanding schemes if it hopes to have its proposals adopted as standards. To get Japan's technological presence in next-generation Internet application technologies, we should offer a large number of prominent technologies and contribute to the standardization of the Grid. Making the most of its Grid projects aimed at developing Grid technology, Japan must promote its technologies in the U.S.-dominated Internet world.

Acknowledgements

The author would like to thank the following people for their valuable comments on the trends in Grid technology, which were given to assist in the preparation of this paper: Dr. Satoshi Sekiguchi, director of the Grid Technology Research Center, the National Institute of Advanced Industrial Science and Technology, Prof. Satoshi Matsuoka of the Global Scientific Information and Computing Center at the Tokyo Institute of Technology, and Prof. Shinji Shimojo of the Cybermedia Center, Osaka University.

Glossary

*1 middleware

Software that functions as an intermediate layer between operating system software and application software. While not including operating system software, middleware provides basic functions that are shared among many applications.

*2 tera

The 12th power of 10, or trillion. A thousand times greater than giga.

*3 e-Science

Applying IT to encourage and accelerate scientific research. It can also indicate the fields of study where IT is intensely exploited to promote research. IT, including the Internet, high-performance computers and the Grid, does not only serve as a research tool but also creates new research field such as bioinformatics. Although "e-Science" is the name of a famous UK project, it often also refers to the concept the project aims at.

References

[1] C. Catlett and L. Smarr, "Metacomputing," Communications of the ACM, Vol. 35, No. 6, pp. 44-52, June 1992.

[2] C. Kesselman, I. Foster, "Globus: Around the World and onto Your Desktop," enVision, Vol. 14.3, NPCI External Relations.

[3] The Grid: Blueprint for a New Computing Infrastructure, Ian Foster and C. Kesselman, ed., Morgan Kaufmann, July 1998.

(Original Japanese version: published in September 2002)