GRID COMPUTING:
A NEW DIMENSION OF THE INTERNET

Wolfgang Gentzsch,
Director Grid Computing,
Sun Microsystems,
Palo Alto, USA

Abstract: The Grid is a distributed computing architecture for accessing Computing, data resources and scientific instruments over the Internet, in much the same way that electricity is delivered over thePower Grid today. It is the next logical step, from Internet, to WorldWide Web, now to technology infrastructure which connects distributed computers, storage devices, mobile devic es, instruments, sensors, data bases, and software applic a-tions, and provides uniform access for the user community. The Grid is the evolving next generation of the Advanced Web, for computing, collaboration and the NSF NCSA National Technology Grid; NetSolve for accessing and sharing mathe matical software; Nimrod for campus-wide resource sharing; SETI@Home for searching for extraterrestrial intelligence; the CERN DataGrid, processing Petabytes of particle data per year (starting in 2006) from its Large Hadron Collider experiment; or the APGrid connecting many computer centers in Asia and the Pacific Rim, in the near future.

Applications for Grids

The list of applications which currently drive the Grid and benefit from communication. And like the Web, the Grid will be ubiquitous. It will simply become the basic IT infrastructure for many businesses and applications, within research, enterprises, and beyond.

Within the last decade, Internet and Web technology and their usage have dramatically improved, as expressed in the exponential laws of Moore, Gilder and Metcalfe, that we soon will see the evolution of the Advanced Web for computing, collaboration and communication. The Advanced Web will be available for any kind of business and private purpose, a cessible for anybody, at any time, from anywhere. There are, however, still a couple of years to go until we simply switch on our appliance and use any information or computing service, available over the Internet or just via wireless access.

Examples of Grids

Examples of current Grids are the NASA Information Power Grid (IPG);
the DoD Distance Computing and Distributed Computing (DisCom 2) Grid; Grid developments is long. It started with large-scale simulations in research, engineering, biology, and drug design. Now the benefits are being seen in the sharing of digital content, application service provisioning, data mining for both research and business applications, collaborative design,

remote usage of experimental instruments, remote medical diagnostics, synthesis of large distributed data sets. And the list goes on and on.

Grids Today

We are ready to build grids today, as mentioned in the above examples. But there is no commonly agreed standard, there is no common architecture yet. Currently, there is manual work involved to set up and run a larger prototype grid for computing, collaboration, and communication.

Since the Grid is still a highly technology-driven software infrastructure, we can identify mainly three user groups: the early adopters, the early majority, and then everybody else. Currently, we are in the early-adopter phase: highly technology oriented users, mostly in research and education and in large engineering companies.

How Easy is it to Build a Grid ?

According to Duncan Ross, director of IT at Cognigen, Corp., their Cluster Grid software has allowed them to experience a 10 to one reduction in hardware costs and double the CPU power. Since implementation, Cognigen scientists now have one extra hour everyday to work on other things. So if main requirements are just high utilization, management, and easy access to the resources in your department, then obviously, complexity is low, and a distributed resource management system will do it.

How Complex Can Grids Be ?

If you want to include resources from different departments, or if you want to do a server consolidation, into one cluster of powerful servers, for all your departments or business units, then obviously, complexity is much higher, and you need a "supervisor" software, like Sun Grid Engine Enterprise Edition, to take care of a fair sharing of all resources, according to company policies.

At the next level, with your enterprise compute and date resources scattered all over the world, a next grid software layer has to be used. Such software is provided by the Globus project or by Avaki, for example, with the additional capability to handle distributed data sets, a single global name space, security, authentication, and authorization.

How to Build a Grid ?

Let's illustrate this with an example: Sun Microsystems' Grid software stack. It consists of Sun ONE Grid Engine for distributed resource management in so-called Cluster Grids, mostly for departmental environments; Sun ONE Grid Engine Enterprise Edition for Enterprise Grids; the HPC Cluster Tools for HPC and parallel applications; the Sun ONE Grid Access Portal for secure, transparent and remote access to the resources; the Solaris Resource and Bandwidth Manager; the Sun ONE Management Center for resource administration; the distributed storage managementsoftware from the newly aquired LSC; Jini and Jxta, for dynamic resource allocation; SunCluster for availability and reliability, and Sun ONE for providing your company's assets as Web Services to your partners and customers. In addition Sun fully supports all the other open grid solutions, e.g Sun ONE Grid Engine is integrated with the major Grid technologies currently developed and deployed in the Grid community, such as Globus, Legion, Punch and Cactus, to form so-called Global Grids.

Examples of Sun Grids

The EPCC Edinburgh Parallel Computing Center is an example for a Sun Center of Excellence in Grid Computing. It is the location of the UK National eScience Center (along with eight regional centers). Basically, the eScience program is to build a UK eScience Grid which interconnects all these distributed computing centers. Edinburgh will evaluate software building blocks like Sun ONE Grid Engine, the Sun ONE Portal Server, Sun ONE Management Center, HPC Cluster Tools, and Sun ONE Studio to build the next generation Grid infrastructure. Through this infrastructure, Edinburgh will deliver compute power to and exchange expertise with its partners in research and industry all over the UK.

Another example is the OSC Ohio Supercomputer Center, another Sun Center of Excellence for High Performance Computing. Together with Sun, OSC is building the Grid infrastructure which enables distributed computing, collaboration, and communication with other partners, e.g. Ohio State University, Universities of Akron and Cincinnati, and Nationwide Insurance.

One more example is the Technical University of Aachen, which is a Sun Center of Excellence for Computational Fluid Dynamics. Among other objectives, the Center will be providing remote access to its TeraFlops Sun supercomputer for researchers on the university campus, and later to other NRW universities, much like an ASP Application Service Provider. Therefore, one of their Grid contributions is the enhancement of Grid Engine toward a Grid Broker, using the software code available in the Grid Engine open source project.

Grids in Industry

Grid computing is already implemented in many industrial settings, on a departmental or enterprise level. For example Caprion Pharmaceuticals has recently implemented Sun Grid Engine for a Proteomics Grid. Grids are also coming into the commercial setting through third parties, since many ASPs and ISPs are implementing grid-enabled access to applications.

Commercial Grids

Unlike technical computing, however, where the key demand is mostly raw system performance and better resource utilization, the demands in commercial environments also cover scalability, manageability, availability, reliability, and transparent access to all resources of an enterprise. Sun ONE Grid Engine Enterprise Edition software, is a key technology to build Enterprise Grids.

A Commercial Enterprise Grid

An Enterprise Grid for a financial institution in Europe, for example, provides portfolio and risk management services to its wealthy customers and investors. In the past, different departments have developed different, market-specific solutions, each within their own home-grown departmental HPC environment. From an enterprise perspective, this IT infrastructure is very inefficient. There is replication, system administration, maintenance, tuning, training, etcetera, in every department. The new Enterprise Grid solution provides economies of scale, access to one common HPC service for all departments, reliability and qualities of service, reduced hardware and software costs, reduced operational cost, and increased productivity.

In most cases, an Enterprise Grid can be built out in two phases. First, we optimize the already existing local environments within the departments. This may take a couple of weeks. No new hardware at this stage. Then, we build an additional central service for more generic tasks, based on the concept of server consolidation, and optimize the overall enterprise environment, through central management.

Grid, The Platform for Commercial Computing

There are mainly four driving forces which finally will lead to the convergence of technical and commercial computing. Number one is the increasing complexity of the applications running on grids. There are many more processes than just technical simulations involved in building products. Computer aided design and simulations are only one part of the overall design, development, production, marketing and sales processes.

The second driver will be the enterprise grid technology. Sun ONE Grid Engine, Enterprise Edition, for example, enables customers to combine all compute and data resources of the whole company, into one large Enterprise Grid. Usually, these resources are mostly for commercial use, only a few of them are used for technical computing. Thus, in an enterprise grid, you can be much more flexible in using and sharing resources. And Sun ONE Grid Engine Enterprise Edition, with its policy module, helps you to remove the "mental firewall" between the different, sometimes competing, teams and individuals in the company.

The third driver is the current trend to enhance grid computing an data services with Web Wervices. The new assets developed in a grid, will be made available through the Sun ONE Web Services environment to your customers, or your partners, or your suppliers. Those assets, created in a grid and offered through a web service environment, will enhance your business in the near future.

Certainly, the fourth driver of grid technology penetrating into the commercial enterprises will be the commercial enterprises themselves. As we recognize the huge benefits grids are providing today in more technical markets, we will investigate how benefits like productivity increases, shorter time-to-market, and increased quality of service, can be transferred into the commercial businesses.

Grids and the Net Effect

The Net Effect now forces us to reinvent the network itself: the data centers, the clients, the applications, and the services. Everything is changing. Everything will be seen in the context of Grids – Campus Grids, Enterprise Grids, Research Grids, Entertainment Grids, Community Grids, Healthcare Grids, Global Grids, and many more. The
network will be service driven, the clients will be light-weight appliances with secure Internet or wireless access to any kind of resources. Data centers will be extremely safe, reliable, virtually always available, from anywhere, on any device. Applications will be part of a wide spectrum of services delivered over the network, such as compute cycles, tools for data processing, accounting and monitoring, with customized consulting, with additional information and communication tools, and with software which allow you to sell or trade your results, over the Internet. Sun ONE Web Services software is one example of such Web Services tools operating on these Grids.

With applications added on top of The Grid, you will be able to build any kind of computing and information service. Such applications are, for example, utility computing, or application service provisioning, or peer- to-peer computing, or complex, federated Web Services. That's where the newly proposed Open Grid Services Architecture (OGSA) will become very important.

Grid Computing as a Utility

What is a utility? Think of water, or electricity, or telephone. For any utility, we can identify three main characteristics. First, you get it when you need it (real time). Second, you don't care about the infrastructure (how you get it). And third, you pay as y ou go, that is to say, you get a bill every month for exactly what you used.

Now, as an example, let's have a look at those over 5,500 Cluster Grids which are powered by Sun ONE Grid Engine software, today. The user submits a job to the Cluster Grid, the Grid Engine Master identifies the best suited and least loaded system, and the job is executed on that system immediately, thus satisfying criteria number one of a utility. And, Sun ONE Grid Engine software provides a single system image for whatever resources are sitting underneath. The underlying hardware, middleware and software are completely transparent to the user, thus fulfilling requirement number two of a utility. And finally, all important consumer information are stored in a file, information about the resources used, the time each resource has been involved, the project for which resources were used, time and resource priotities, etc. With this data, the user can build his own usage report, on a daily, weekly, or monthly basis, and use it for accounting and billing. Here we are, with requirement number three.
Thus, Grid technologies today encourage utility computing at the departmental and enterprise level.

What Comes After The Grid?

There will be no real disruptive, immediate change in the IT technology or in its usage, for many good reasons. We have so many new technologies on the table that we will be busy enough to efficiently implement and deploy them. This takes time. Then, after we have interconnected all kind computing devices through Grid infrastructures, the next step will be to embed and hide these devices into any kind of thing that serves our daily business and private needs, in our houses, cars, airplanes, clothes, the environment, and some even in our body.

We are still far away from the Grid infrastructure we envision today. We need to agree on one common unifying architecture for the Grids, and on standards and protocols for connecting all the Grid building blocks. Today, hundreds of computer scientists in the Global Grid Forum standardization working groups are concentrating on these most important tasks. We still have to develop much better software for managing, monitoring, billing and trading of any kind of computing services. And we have to make the computing environments absolutely user-friendly, fully secure and reliable.

How The Grid Will Impact Our Future ?

Let me answer with an analogy. What singular event in the 19th century had the greatest impact on our community? Answer: Among others, certainly the steam engine, which got us into the Industrial Age. The steam engine enabled mass production and shortened production cycles dramatically. The steam engine was a time machine.
So, what was one of the most important inventions which changed our lives in the 20th century? Answer: the combustion engine. The combustion engine, too, is a time machine. Put it in a car, in a plane, it enables us to compress time to get from A to B much faster.
Another time machine.

Like this, the Grid Engine will be the next time machine. The Grid Engine will shorten time to market for many products dramatically, enlarging computing space (or cyber-space) seamlessly, bringing information, knowledge, education, and healthcare to the masses, and thus compressing the time to design, simulate, test, build and sell our products of the future.

The Grid is becoming the time machine of the 21st century.