東海大學資訊工程與科學系

碩士論文

指導教授：楊朝棟 博士

應用網格架構之數位學習平台

Implementation of Grid Architecture for e-Learning Platform

研究生：何信權

中華民國 九十三 年 六 月

應用網格架構之數位學習平台

研究生：何信權　　　　　　　　　　　指導教授: 楊朝棟　博士

東 海 大 學 理 學 院

資 訊 工 程 與 科 學 系

摘要

隨著網際網路的快速發展，知識的傳遞速度已不可同日而語，影響所及，傳統的學校教育方式也興起了重大變革，例如能夠突破時間與空間限制的數位學習(e-Learning)近年來也成為普遍的教學方式之一。目前校園內大部分的數位學習平台架構，都是以單一電腦或伺服器為架設基礎考量，一旦工作量增加時只能不斷更新或升級軟硬體，對於經費不充裕的學校而言是一大負擔。因此本研究的主要目的是利用「網格運算」(Grid Computing)技術，將學校內的閒置電腦加以整合成為數位學習的教學平台，以取代購置昂貴的高階伺服器。其他擁有相同平台的學校，也可以藉由網格架構互相分享資源。如此加入網格的學校或組織不但可以充分運用自己的閒置電腦資源，隨著網格組織的擴大，同時還能源源不絕分享其他學校或組織的電腦資源及運算能力，以真正達到資源共享互蒙其利的目的。我們也希望藉由此研究將建立網格數位學習平台的經驗及技術提供給需要的學校或單位，以節省其建置的時間和成本。

Implementation of Grid Architecture for e-Learning Platform

Student: Hsin-Chuan Ho Advisor: Dr. Chao-Tung Yang

Department of Computer Science and Information Engineering
Tunghai University
Taichung, 407, Taiwan, Republic of China

Abstract

The fast development of Internet has accelerated the speed of transmission of knowledge which is beyond that in the past. Also, by the influence of Internet, traditional academic education has been in revolution, such as e-Learning breaking the limitations of time and space, has become the general educational methods. Currently on campus, most of e-Leaning platforms are constructed on the basic consideration of either single computer or server. Once the workload increases, renewing or upgrading hardware has become the only solution, which is a great burden to a school with insufficient budget. Therefore, by using Grid Computing technology, this study aims to integrate those unrelated computers in schools to replace high level server as the teaching platforms of e-Learning. Thus, not only the computer sources in schools can be fully used and applied, and moreover, constantly sharing the computer sources from other schools or associations as well as the computing capacities. We also expect the outcome and experiences of this research can provide reference to the schools that wish to develop e-Learning environment for saving the cost and time in developing a similar system.

Acknowledgements

I would like to thank all of those people who have helped me through the completion of this thesis. In particular, I would like to thank my advisor, Dr. Chao-Tung Yang, who introduced me to this topic and gave me a broad support and guidance. I would also like to thank Professor William C. Chu, Dr Shian-Jeng Yu and Professor Yi-Min Wang for their valuable comments and advice on my oral defense for this thesis.

There are many people whom I would like to thank, especially to my wife and my parents. Without their encouragement and support, I couldn’t pursue further education in the graduate school of Tunghai. Many of my classmates have encouraged and supported me through this research. For this reason, I could complete my work without fear.

Lastly, I would like to thank all my coworkers and friends whose unconditional support making this thesis possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Compared to the conventional learning method, e-Learning has overthrown the former method of teaching and class attending. Regardless of teaching personnel or learner, one can more flexibly propagate and absorb knowledge. Nowadays, either enterprises or schools are all largely adopting e-Learning as their method of training or instruction [1].

Generally, e-Learning Platform requires high capacity of Storage Device to store contents of teaching such as Disk Array, NAS (Network Attached Storage) or SAN (Storage Area Network). To schools with insufficient budgets, this expensive equipment is often not affordable. However, as long as we can examine our present equipment, we can discover that many available idle computer resources can be utilized. For example, a personal computer with 20GB HDD that was bought two years ago, if this computer is only used for browsing online and word processing, the space of its hard disk only requires 30% of the total capacity, which is 6GB. The remaining 14GB will become idle. Therefore, we can connect some computers that are within the dispersed facility. It is used for teaching or administration and then their idle storage spaces can be combined together by the means of Grid Computing technology. In this way, it is possible that the storage space obtained will not be less than of its expensive storage device. This research will work on the connection of the idle storage devices in the academy with the techniques of Data Grid to substitute an expensive storage device such as Disk Array or NAS and act as an e-Learning platform storage device.

1.2 Contribution

This research hopes to integrate those idle computer facilities in academies by Data Grid techniques so as to save cost and make best use of resources. A huge amount of software and hardware equipment purchasing expenditure can be saved and the goal of complementation and sharing of resources among the schools can be achieved.

For those academies that have insufficient budges, they can then obtain better services and enormous teaching resources through the techniques of Grid. With the continual participation of partner academies, the expected accomplishments and experiences of this research can be provided as reference to those academies which hope to develop E-learning so as to save their cost and time to develop the similar system.

This research also aims to set up a method of automated adjustment of grid portal framework through web mining technologies. It will make the framework of Grid Portal adjustable on its own by following the browser's preference and to make the web framework meet actual needs of the browsers.

1.3 Thesis Organization

This thesis begins with a brief introduction toward the motivation of this research. In chapter 2, we discuss the background of e-Learning, Grid computing technologies, web services. And we consider the need to apply Grid computing technologies to e-Learning. In chapter 3, we present the architecture for Grid-based e-Learning platform. In chapter 4, we show the grid Portal structure optimization method using web mining algorithm. Related work is introduced in the following chapter. Finally we make a brief conclusion and future work in chapter 6.

Chapter 2

Background

2.1 e-Learning Technologies

Due to overall popularity of the Internet, e-Learning has become a fervent method of learning in recent years. Through the use of Internet, learners can freely absorb new knowledge without the restriction of time and place. Many companies have adopted e-Learning to train their employees. The e-Learning system can make an enterprise more competitive by increasing the knowledge of its employees. As we know, e-Learning has become one of the most potential e-commerce businesses of all time.

With the development in communication and network technology in recent years, under the gradual improvement of network bandwidth and quality, the real-time transmission of high-quality video and audio is now a reality. Therefore, the transmission of multimedia and relative network application technologies have gradually been developed and become popular such as the technology of Distance Education, Video Conference and Video on Demand etc.

Distance Education is a very essential link in the application of e-Learning. Regardless of teaching courses in a campus or the training courses for enterprises, Distance Education is a very effective method of learning. With conventional teaching method, learners must gather the same time and same location for attending classes, which the enterprises have numerous learners [2]. It frequently causes exorbitant was expenditure of money and time. The advantage of Distance Education is it can cope with the obstacle of geographical location. Making students at remote site feel more like being in an environment of attending classes in a regular classroom. Furthermore, it can conserve their time, money and energy for their transportation between classes.

Although Distance Education has many advantages but the biggest obstacle is the investment of the facilities. In order to facilitate more enterprises and schools to apply Distance Education, this research has combined Grid Computing technologies for building one of the set for e-Learning platform that consist expandability and resources sharing capability.

The development of current Distance Education System for e-Learning can roughly be classified into three types as follows

2.1.1 Multicast Teaching System

This type of system can enable the teacher and students to be located in different places. By using network technology, the video/audio classroom and multimedia teaching materials can be transmitted in real-time to the remote site classrooms. Furthermore, it allows two-way real-time communication between the teacher and the students in remote site classrooms.

2.1.2 Virtual Classroom Teaching System

This type of system adopts one set of teaching management system to simulate the scenario of attending class in a classroom (such as teacher's lecture, holding examination, specified assignment or question and answer, student proposed question or participating examination, etc). Teacher and students can be linked to the teaching management system at any time through the Internet.

2.1.3 Video on Demand Teaching System

This type of system adopts a technology called VOD (Video on Demand). Students can obtain the teaching-learning materials through the Internet by using a computer or a television which is furnished with Set - Top Box. This will be able to process the distance learning in the accordance with a personal learning speed by controlling the broadcasting process.

The teaching mode of this research combines with the Multicast and VOD teaching modes. Real-time video teaching is performed at a fixed time but the rest of the time is only for web teaching. For instance, the video teaching course is produced as a video stream file for broadcasting online after every class for the convenience of the learners who could not watch the real-time video teaching course.

2.2 Grid Infrastructure

Grid infrastructure supports the sharing and coordinated use for resources in dynamic global heterogeneous distributed environments. It involves resources that can manage computers, data, telecommunication, network facilities and software applications provided by different organizations [3]. Namely, a Grid is a collection of distributed computing resources available over a local or wide area network that appears to an end user or application as one large virtual computing system.

2.2.1 The Concept of Grid Computing

Distance Education or e-Learning can only be put into practice by relying on computer technology. Nevertheless, the structure of traditional computer's computing, including software, hardware and Internet, etc. will be obstructed because of an increasingly workload on every other phase of time. That will lead to a slower speed of computing. In order to solve this problem, it is necessary to upgrade or renew the software and hardware, which has often become very difficult for the academies that have inefficient budgets.

Thus, we employ a technique called “Grid Computing” to solve this problem. “Grid Computing” can connect the scattered computer resources through the links of the Internet together with the assistance of a proper operation of system and software. These resources include computer facilities, storage and various kinds of input and output devices.

When we turn on the PCs or facilities on the grid, other computers among the grid will search for them to begin processing and computing the relative work. This will not only uses the resources of local computer, but also obtains the powerful virtual resources from other remote computers, including the computing and storage capability, application programs, database, I/O device that are spread all over the whole grid. This is then called “Grid Computing”. In other words, “Grid Computing” can be considered as a virtual super computer [4].

For example, if the host needs to finish a task in an hour, the technique of “Grid Computing” can search for another idle computer. The Grid further transmits the task to these computers to be able to perform an accordance of the proper allocation and then send back in return. Due to the efficient usage of computer resources, this task will be finished in less than ten minutes. These computers can be in different organizations or locations, even in different countries. They could just share the computer resources with one another via Internet [5].

Nowadays, many academies could not renew their computer software and hardware facilities due to lack of budget. However, if we examine the present facilities in the campus, we can find there are many idle computers which can be used. For example, if a PC containing Pentium III CPU, 128MB RAM, 20GB HDD, which was purchased a couple of years ago that is only used for browsing online or word processing, its computing capability and storage spaces are absolutely enough. But if it is used as a server, it would then be much inadequate. We can connect some of those PCs as above spread over for on academy for teaching or administration. The connection of their computing capability will not be lower than of an expensive high speed server. Furthermore, those academies that could not replenish their computers' software and hardware facilities could share with other academies' resources by the technique of "Grid Computing".

Although many academies have now adopted e-Learning for teaching, the software and hardware facilities that the academies use are very different. It has become a cause of major difficulties in sharing the teaching resources with one another. The technique of Web Services can integrate the different information system within the Grid to solve the above mentioned problem. Therefore, the concept of this research is based on integrating Grid Computing with Web Services to build up an e-Learning platform [6]. Figure 2.1 shows the concept of grid computing.

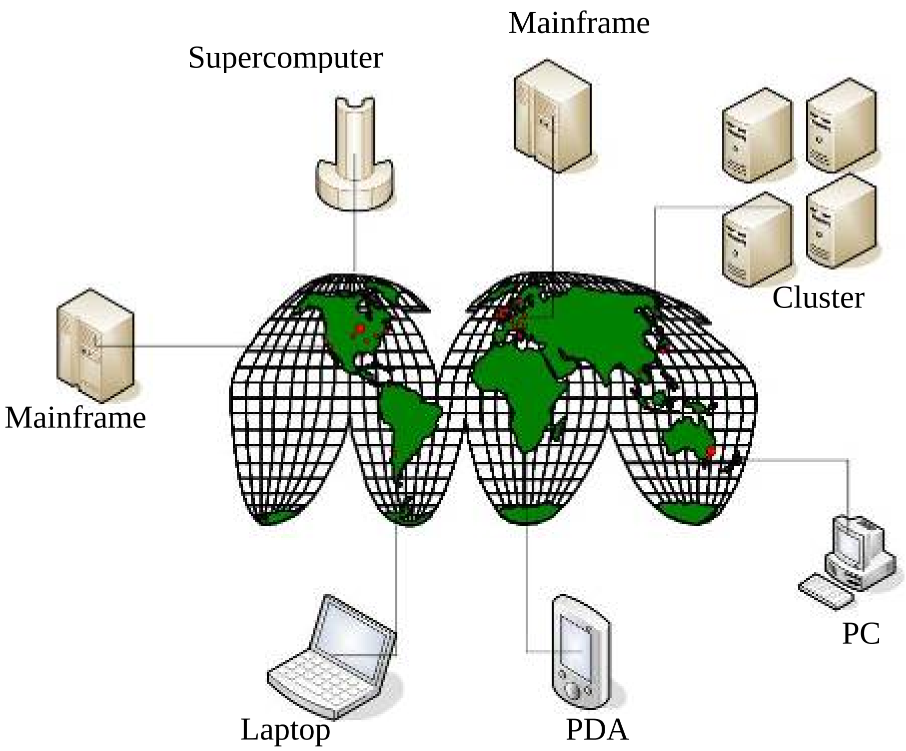

Figure 2.1: The Grid computing infrastructure

2.2.2 Globus Toolkit

The Grid represents the common properties that have very large, distributed, dynamic and cross the boundaries of human organizations [7]. Due to these characteristics, its architecture must be very complex and its relationship is also like mazy. The critical point to realize Grid is to establish a common open standard. The base of its core is an open set of standards and protocols - e.g., Open Grid Services Architecture (OGSA) - that enable communication across heterogeneous, geographically dispersed environment. With Grid computing, organizations can optimize computing and data resources, pool them for large capacity of workloads, share them across networks and enable collaboration [8]. Some protocols and applications have been proposed and implemented the actual work, such as Globus Toolkit [9], DICE and the NASA's information Power Grid.

The Globus Toolkit is an open architecture, open source software of services and software libraries that support Grids and Grid applications. It facilitates the creation of computational Grids. It provides software tools that enable coupling with people, computers, databases and instruments. With Globus, you can run job on two or more high-performance machines at the same time, even though these machines might be located far apart and owned by different organizations. Globus software helps scientists to deal with a very large scale of datasets and complex remote collaborations. The toolkit includes software for security, information infrastructure, resource management, data management, communication, fault detection, and portability.

The toolkit components are most relevant to OGSA that are the Grid Resource Allocation and Management (GRAM) protocol, the Meta Directory Service (MDS-2) and the Grid Security Infrastructure (GSI). These components provide the essential elements of a service-oriented architecture.

- GRAM protocol: provides for secure, reliable, service creation and management of arbitrary computations.

- MDS-2: provides a uniform framework for information discovering through soft state registration, data modeling and a local registry.
- GSI protocol: supports single sin on, authentication, communication protection and certification mapping.

2.2.3 Data Grid

Some newly developed scientific studies such as high energy physics and computational genomics demand access of large amounts of data. A requirement that leads to such considerations as data file management, replicated files management and transfer, and distributed data access management. The Data Grid infrastructure is meant to integrate the data storage devices and data management service into the grid environment. Data grid consists of scattered computing and storage resources, which, though located in different countries, remain accessible to users [10].

In this research, we have adopted Globus Toolkit as infrastructure for Data Grid. Globus Toolkit provides solutions to such considerations as security, resource management, data management, and information service. A large number of research projects, such as GriPhyN [11], PPDG [12] and EU DataGrid [13], are based on Globus Toolkit.

Globus Data Grid comes in two layers. In the Low Level are Data Grid Core Services, and the upper layer are the High Level Components [14]. Figure 2.2 shows the Data Grid Architecture.

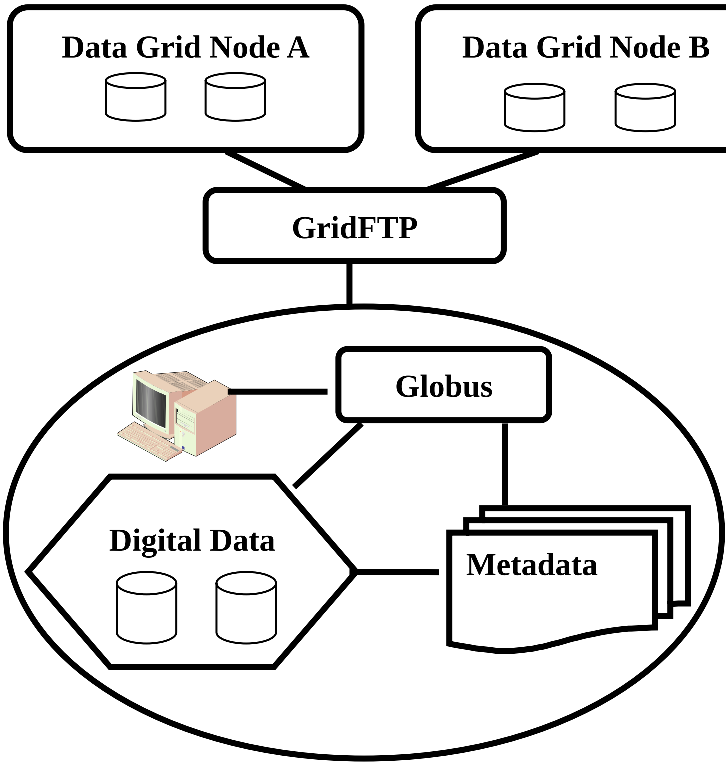

Figure 2.2: The Data Grid Architecture

The storage system is a basic data grid component. It defined and covers all of the storage technologies capable of adding, deleting, reading, writing and operating file instances [15], such as HPSS, DPSS (Distributed Parallel Storage System). Data access service is set up as a mechanism for accessing, managing and transferring data in the storage system [16].

Metadata service is a means of managing and accessing Metadata, which contains Data Grid information. Application Metadata includes information describing files and information on data environments. Replica Metadata is applied to manage replication of data objects, and it is a way to help map logic files with physical files.

Resource Management are responsible for storage system, networks and other data grid resources to assure end-to-end efficiency, technical assessment of efficiency test, as well as crucial resources. Grid Security Infrastructure (GSI) [17] provides environment authorization and certification mechanism to a large number of users.

Under the framework of Data Grid lies a key element that is known as replica management [18], which is important to successful processing of large amount of data in Data Grid. It is mainly to decide when and where to set up replica and it provides information on the location of Replica, including the key functions given as follows:

- Registration: To register new files in Replica Catalog, where the new files are accessible to users.
- Creation and deletion: To create or delete registered replicas.
- Publishing: To reproduce unpublished files onto destination storage system for publication.
- Copying: To reproduce files among storage systems.
- Query: To check physical location of the storage system where replicas are stored.

Data Grid Applications will produce a huge data, and managing replicated data in Data Grid is a major problem. Replica Location Service (RLS) is one component of Data Grid management architecture. An RLS service provides a mechanism for registering the data of replicas on the grid and discovering them. The RLS architecture is based on following five elements:

- Local Replica Catalogs (LRCs) that contain mappings from logical file names (LFNs) to Physical file names (PFNs).
- Replica Location Indexes (RLIs) that aggregate state information about one or more LRCs and supports discovery of replicas at multiple sites
- Using soft state protocols such as LDAP to maintain RLI state
- Optional compression of soft state updates reduces communication, CPU and storage overheads.
- Manage RLS membership for location of LRCs and RLIs.

2.2.4 Grid Portal

A grid portal allows communication between the outside world and the grid itself, and a grid portal is always of huge and complex framework. So far, it is mainly applied in AP portal (Application Portal) and User portal. An AP portal [19] allows specific grid operation for specific applications, such as Astrophysics Simulation Collaboratory (ASC portal) [20] and Diesel Combustion Collaboratory [21]. User portal provides special services to specific public or researchers, such as HotPage Portal user portal [22], the Gateway project [23], and UNICORE [24].

There are several technologies in the construction of the grid portal. One is **GPDK (Grid Portal Development Kit)** [25], which have the Tomcat Application Server as its development platform through Java technology, including Java beans, Servlets and JSPs. In the communication with Java CoG toolkit and Globus toolkit, Globus toolkit is a grid middleware, which is responsible for grid resources allocation, task coordination and security; this makes the soul of grid operation. In addition to Globus, Sun Grid Engine [26] of Sun Microsystems is another Grid Middleware.

Another grid portal technology is known as **GridPort** [27], which is a toolkit successfully developed by SDSC (San Diego Supercomputer Center) with its NPACI [28] project in 1999. The purpose of this technology is to allow the computer experts involved in the NPACI project to be accessible to the grid resources, account management, mass data transfer and task allocation through a web-based interface. GridPort is a multi-tier framework made up of Client tier, Portal tier, Portal Services tier, Grid tier, and Resources tier.

The other grid portal is **OGCE** [29] **(**Open Grid Computing Environments Collaboratory) Portal which developed by NMI (NSF Middleware Initiative). This portal technology provides access to Grid technologies through sharable and reusable components for web-based access to scientific and business-oriented applications.

The primary requirements of Grid portal system [30] from a user's point of view must include :

1. **Security:** Users will visit a portal using a web browser and they will authenticate by means of a user-id and password. While better technologies exist for authentication, this is one of the demands of the user. More secure systems, such as smart cards are possible, but unlikely to be deployed anytime soon.
2. **Remote file:** Tools to access to file metadata directories and remote file archives are a central requirement for the portal. Simple tools for Grid FTP are essential, but many files are likely to be managed by a virtual data system, where data is cataloged, curetted and staged by back-end grid services.
3. **Remote job:** The ability to submit jobs to the Grid for execution and monitoring is a standard requirement for portals. Users with allocation on specific resources

want to be able to see the job queues on those resources and consult scheduling assistants. They also need to be able to keep track of job execution and understand when things fail by reading logs.

4. **User's information services:** Access to directories and index tools is an essential role of the portal. Each user should have a private, persistent store of references to important information that they have stored on the Grid.
5. **Application interfaces:** The key to scientific portals are being able to hide Grid details behind the useful application interfaces. The user needs to be able to launch, configure and control remote applications in the same way a user uses a desktop applications.
6. **Access to collaboration:** For any Grid Organization, it must be possible to share resources. This includes the ability to use real-time collaboration tools as well as asynchronous collaboration.

In grid portal design, portal data that needs to be described using XML. Although grid computing can integrate the computer resources inside the grid organization, these resources can be located on different kinds of platform. The basis of information transmission has to reply on XML Web services. In fact, Web Services is already a standard structure of Globus Toolkit 3.0 called OGSA [31][32] (Open Grid Services Architecture), which is also called "Grid Service".

On the above Grid portal technology, all grid portals are established to provide grid users with a web-based interface to use for the resources in the grid. Successful adjustment of the grid portal framework is by means of Web Mining, it will surely bring about more efficient resource access.

2.2.5 Web Services

In the Grid environment, the critical point on how to combine the different heterogeneous resources. Web services defines a technique for describing software components to be accessed via internet, it is a communication media between different platforms. Web services standards are defined within the W3C that has the most industries support, and interact between the components in the service processes that are based on XML, SOAP, WSDL and UDDI. Web services architecture is shown in Figure 2.3 [33].

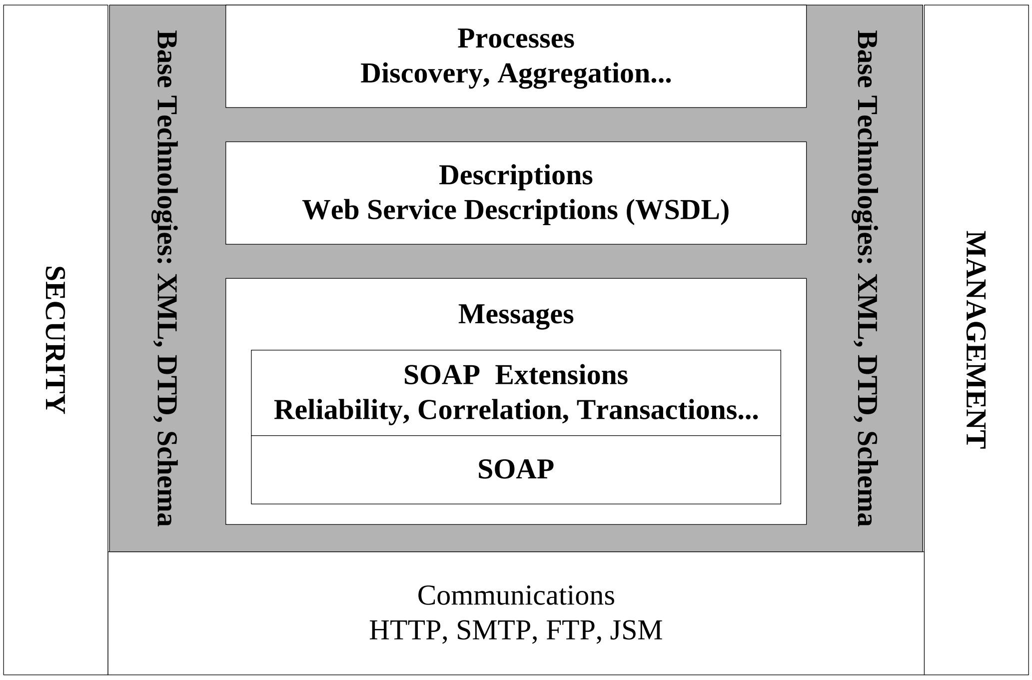

Figure 2.3: Web services diagram

The term Web services is described by an XML that covers all the details that is necessary to interact within the services, including message formats, transport protocols and location. SOAP provides a means of messaging between a service provider and a service requestor. It is independent of underlying the transport protocol. SOAP can carry on HTTP, FTP and SMTP. WSDL is an XML-based language to describe a Web services how to access them, it provides a formal framework to describe services in terms of protocols, servers, ports and operations that can be invoked. The specification that provides a SOAP binding which is the most natural technology to be used for implements a Web Services. Universal Description, Discovery and Integration (UDDI), apparently, it provides the registry and search mechanism for Web services. Concisely, WSDL describes the format SOAP messages, and UDDI serves as a discovery service for the WSDL descriptions.

The Grid emphasizes the usage of Web services does not only indicate the use of SOAP for all communications. If needed, alternate transport can be utilized, for example to achieve higher performance or to be able to collide with a specialized network protocols.

In the Grid environment, the critical point is on how to combine the different heterogeneous resources. XML-based metadata is a popular problem solving, so it has been widely used. The XML document can not only help manage facilities, but also interchange between different databases. The interface based on Web services that can integrate not only the web resources easily, but also make the occurrence much faster to duplicate. If the results of this research are to be duplicated to other academies, the process of duplicating will become much easier that it is based on the open structure of web service.

Chapter 3

e-Learning Platform with Grid

3.1 Platform Architecture

3.1.1 Hardware and Software Configuration

As we know, the prices of PC keep falling down, but the prices of high-end servers and other supercomputers that are still remain expensive. Therefore it has become a costly productivity, i.e., similar computing task can be executed by the PC Cluster system that has a lower cost than others. In applying a high-end computer and for the difference of efficiency is not enormous. Therefore, using PC Cluster system to substitute as a high-end server makes the cost outcome and efficiency more possible.

Our Grid architecture is implemented on top of the Globus Toolkit, called grid-cluster. It is built under the three Clusters to form a multiple cluster environment. Cluster1 and Cluster2 consist of four PCs that each has a master node and the rest are slave nodes. Cluster3 also consists of four PCs, which have a master node and three slave nodes, as it is shown in Table 2.1. Each node is interconnected through 10/100 Fast Ethernet Card to a Switch HUB. At present, the Education Grid includes three database nodes, with two nodes located at Tunghai University and another node that is situated at Lizen High School.

The operating system is RedHat Linux release 9, and Data Grid application that we use Globus Toolkit 3.0. In general, basic Grid technology services include security service, scheduling, data service, database service, user service, application management service, autonomy and monitoring service, information service, composition service and message service.

Table 2.1: Hardware configuration of Grid-Nodes

Node1*				
No	Domain Name	Host Name	Processor (CPU=4)	Memory
1	beta1.hpc.csie.thu.edu.tw	Beta1*	Intel Celeron 1.7GHz	512MB
2	beta2.hpc.csie.thu.edu.tw	Beta2	Intel Celeron 1.7GHz	256MB
3	beta3.hpc.csie.thu.edu.tw	Beta3	Intel Celeron 1.7GHz	256MB
4	beta4.hpc.csie.thu.edu.tw	Beta4	Intel Celeron 1.7GHz	256MB

Node2				
No	Domain Name	Host Name	Processor (CPU=8)	Memory
1	gamma1.hpc.csie.thu.edu.tw	Gamma1*	Pentium III × 2	512MB
2	gamma2.hpc.csie.thu.edu.tw	Gamma2	Pentium III × 2	512MB
3	gamma3.hpc.csie.thu.edu.tw	Gamma3	Pentium III × 2	512MB
4	gamma4.hpc.csie.thu.edu.tw	Gamma4	Pentium III × 2	512MB

Node3				
No	Domain Name	Host Name	Processor (CPU=4)	Memory
1	lz01.lzsh.tcc.edu.tw	lz01	Pentium IV × 1	256MB
2	lz02.lzsh.tcc.edu.tw	lz02	Pentium IV × 1	256MB
3	lz03.lzsh.tcc.edu.tw	lz03	Pentium IV × 1	256MB
4	lz04.lzsh.tcc.edu.tw	lz04	Pentium IV × 1	256MB

* stand for Master node of the cluster, the others is slave node

3.1.2 The Front End – Grid Portal

By using Access Grid technologies, the school can integrate their training courses and materials into Grid environment which provide more flexible teaching mode.

At first, it is necessary to set up a Portal Web Site in the Grid system; it does not only provide services to the members within the Grid, but also acts as a teaching platform for the other academies through the Internet. Those organizations or individuals who have not joined the Grid will be restricted to use some of the accessible resources.

Our Education Grid utilizes NMI's (NSF Middleware Initiative) OGCE Portal [34] as the Grid Portal Site for the network. The OGCE Portal includes the following functions:

1. Utilizing the Grid Portal Site to post in the Discussions board and communicate with friends inside the Chat rooms, at the same time providing the users with the latest grid-related technological updates in the News section.
2. Monitoring resources on the Grid, such as the operational condition between each node.
3. Utilizing the Grid Portal Site to submit task for operation in the Grid.
4. Using GridFTP to transmit data within the Grid.
5. Using the Proxy Manager to manage the CA in the Grid.

The OGCE Portal framework provides a general portal architecture that supports the virtual organizations comprised of scientists and project developers. And also the API for the development of reusable, modular components that serve to access the services being developed within the Grid organization.

OGCE Portal required tools and application that are listed below:

- Java JDK >= 1.3
- Jakarta Tomcat >= 4.1.18
- Jakarta Ant >= 1.3.5.1
- MySQL JDBC Connector >= 2.0.14

3.1.3 The Back End – Contents of Data Grid

The experimental location of this research is at Lizen High School in Taichung, Taiwan. At the present, the school has already possessed powerful English CAI application software. An interactive English training course for speaking, listening, reading and writing can be processed in all directions. The real-time testing is implemented during the class for evaluating as a result. For establishing this platform, we shall design a system that can be in accordance with the architecture as shown in Figure 3.1.

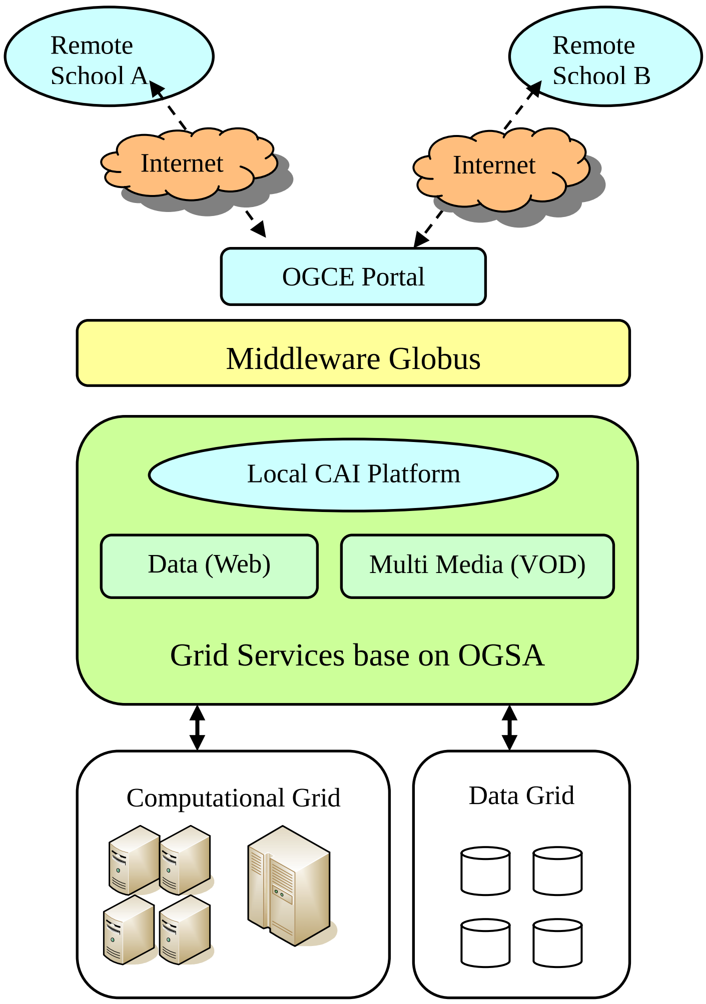

Figure 3.1: e-Learning System Framework

The above-mentioned teaching software, along with the growing number of users and increasingly teaching materials, the loading of the server will continuously increase. Therefore, the storage device needs to be updated or transformed once in a certain period. To overcome this problem, this research will work on the connection of the idle storage devices in academy with techniques of Data Grid to substitute the expensive storage device such as Disk Array or NAS and act as an e-Learning platform storage device.

To simulate data server and application server by using PC Cluster and LAN, which is the front end of the basic computational capability and data source for e-Learning once it has been obtained. Currently to integrate the teaching resources including English CAI training software and database is expected. Besides the above mentioned CAI training software, there is also a VOD (Video on Demand) system, which can provide services for multi-media stream service connection via Web, which can integrated with CAI training software.

3.1.4 System Implementation

Our Education Grid utilizes a DataGrid structure as the implementation technology. Within the DataGrid, the RLS (Replica Location Service) [35] plays a vital role providing a comparison feature between logical and physical file names. At the same time, DataGrid users are given a convenient interface for saving and loading resources. It alleviates the worries of where and how users' resources are stored.

In the boundaries of the Globus structure, the RLS Server is separated into two parts. First, the Local Replica Catalogs allow the LFN to correspond against the PFN on the physical storage system. Another part is that the RLI (Replica Location Indices), corresponds to the LFN against the LRC. Although the LRC and the RLI possess different functions, both can be installed on the same server. In case that they are installed on different servers, the LDAP is utilized as the communicative protocol. To allow a streamlined system structure, we will install the LRC and the RLI together and designate them as the most important parts within the DataGrid system functions.

Figure 3.2 shows the complete system structure that we intend to accomplish. The bottommost level is the physical database. We will employ the usage of MySQL Server as the database system. At present, the Education Grid includes three database nodes. In which the two nodes are located at Tunghai University and the other node is situated at Lizen High School.

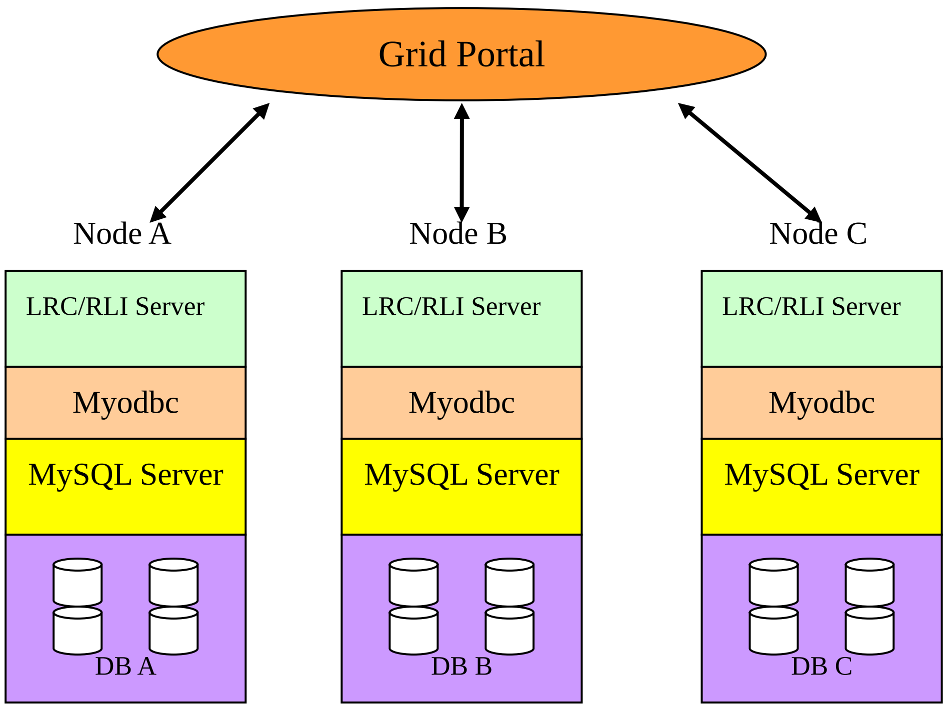

Figure 3.2: e-Learning Portal Framework

Proceeding to the upper level, the database is connected by myodbc and ODBC. Both of which provide the interface for connection with the LRC/RLI Servers. With regard to the copying and moving of files on the DataGrid, the GridFTP Protocol is used as the transmission device. GridFTP is a highly-efficient that allows speedy, safe, and dependable file transfers.

Control of the LRC/RLI Server requires another RLS client program. We installed the RLS client program together with the Grid Portal, thus allowing the network users to access resources on the grid through the Grid Portal.

Figure 3.3 and Figure 3.4 displays a screenshot of the Educational Grid's Grid Portal Site. We plan to place some multimedia educational materials, with different content for each node. For instance, Node A can be used to for math materials, Node

B can be used for English materials and Node C can be use for biological materials. At the same time, backup and query of files can be achieved through the aforementioned RLS Server. We will thus be able to truly realize the establishment of an educational platform.

Figure 3.3: e-Learning Portal Screenshot-1

Figure 3.4: e-Learning Portal Screenshot-2

3.2 Architecture of Partner Schools

The system was established for this research that can apply not only in the school but also to provide for other schools via the Internet for sharing its teaching materials. We hereby propose a Grid virtual organization that has a concept of “Partner School”; i.e., the “Partner School” can be “Resource receiver” that shares the existing resources on the grid, but it can also be a “Resource provider” that provides the resources. Every partner school can provide different teaching resources according to its professional specialty. For instance, School A can provide mathematical category whereas School B can provide literature category for its teaching resources and so on. In such a way, each of the partner schools can concentrate on its resources for developing on its own professional specialty. Meanwhile, it can also be used for professional teaching resources to other schools. The expenditure of teaching cost investment can then be avoided. Figure 3.4 shows the framework of partner academy.

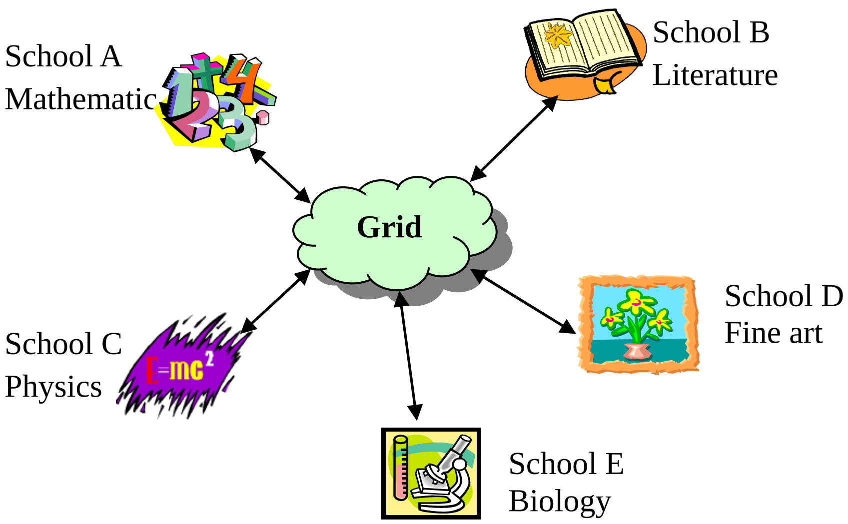

Figure 3.5: Framework of partner academy

Chapter 4
Optimization on Grid Portal Structure

4.1 Web Mining

In a grid organization, the grid portal allows communication between the outside world and the grid. The major function is to provide users with a web-based interface for the grid service and resources. To make the grid portal run more efficient way, we use the web mining technology to adjust its framework. To make it congregate users' needs.

Web mining is a type of application of data mining and it is categorized into "web content mining" and "web usage mining". Web content mining deals with surveying the contents of the web and it is mainly applied to analyzing web contents and reinforcing web-searching capability, such as a search engine. Web usage mining mainly deals with analyzing the web users' browsing behavior through web log, so that the web administrator will become aware of the operation of the website before modifying the system, for the purpose of more efficient operation [36].

4.1.1 Web Usage Mining

Web Usage Mining suggests analysis of user's browsing behavior through the web log. It enables the web administrator to be aware of the operation of the website before modifying the framework, for the purpose of more efficient operation. Its application comes as follows:

(1) Analysis of web flow

In the web log, the title and size of a specific web and a specific access are duly registered and pattern for the web flow. In the [37] study, variation of web flow in each day of the week is registered in the web log. Analysis of it will become more applicable on the adjustment and planning of network bandwidth.

(2) Improvement of web accessibility

By analyzing the web log, we have come to realize what popular websites and files are received more in an access rate. Consequently, we can place those files in the Web Proxy server or cache server. In the [38] and [39] study, we can analyze the web log and make a forecast of the popular web pages to decide on its priority of the pages in Cache.

(3) Analysis or forecast of browsing behavior

This is a common subject being studied in Web Usage Mining. Robert Cooley et. al. [40][41] is one of the few experts that introduced the idea. Other studies aim to improve or propose new algorithmic methods [42][43] for more efficient analysis. On the other hand, the web mining tool WUM [44][45] used in this study is theoretically based on aggregate tree, and MINT language that resembles SQL is used for analysis of the web before reaching to the patterns that we are interested in.

4.1.2 WUM Algorithm

WUM is the system tool that is use for this study [46]. WUM is theoretically based on Aggregate Tree. Suppose a web involves a browsing session of 5 browsers, and a, b, c, d, e, f, h are codes for the web pages. Table 4.1 set as the following browsing record.

Table 4.1: WUM Aggregate Table

No	Session	Sequence	Appearances
1	ab	(a,1)(b,1)	40
2	ac	(a,1)(c,1)	20
3	abcde	(a,1)(b,1)(c,1)(d,1)(e,1)	30
4	bcbf	(b,1)(c,1)(b,2)(f,1)	5
5	abdfhe	(a,1)(b,1)(d,1)(f,1)(h,1)(e,1)	10

For example, In the Session 1—ab, the 1 in (a,1) stands for the sequence of the page for the browsing session. 40 suggest that 40 sessions took place in the web log. Now in figure 4.1 set as the following Aggregate Tree.

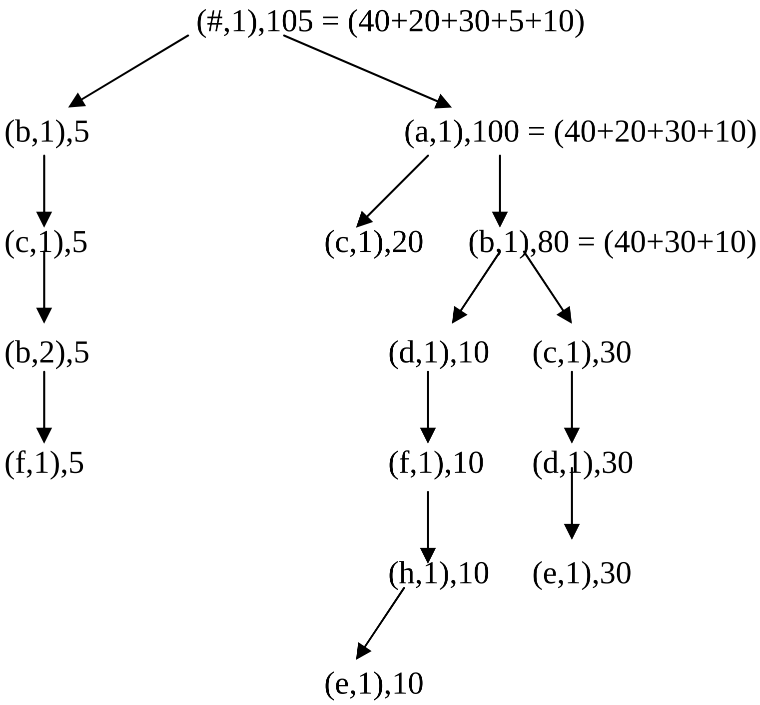

Figure 4.1: WUM Aggregate Table

Given the Aggregate Tree above, suppose we are to locate routes meeting the following requirements:

(1) Appearance rate of the first page X at above 85%
(2) Appearance rate of pages X through Y equals 80%
(3) Appearance rate of pages X through Z at above or equals 40%
(Pages X, Y and Z stand for the pages in the web and they could be any of a, b, c,

d, e, f or h). Then we could use WUM MINT Query commands to locate the best route.

```
SELECT T
FROM NODE AS X, Y, Z
        Template X*Y*Z AS T
WHERE X.Support >=85
AND (Y.Support / X.Support) >= 0.8
AND (Z.Support / X.Support) >= 0.4
```

For instance, an administrator of Bioinformatics Grid Portal is able to locate the routes and pages in the session of browsing from DNA Sequence through the RNA Sequence and then to Expression Database pages that meet the requirements. To figure out which of them are the most popular ones, a text will then be built with MINT commands which allow the user to analyze the data obtained.

4.1.3 Greedy Algorithm

Data obtained through WUM analysis can be taken as a pattern, which comes from a kind of a rule. An example is that the browse from page A, page B to page C must follow a number of rules. To locate the most beneficial one, other methods are needed. In this case, we opt for the Greedy Algorithm.

Greedy Algorithm is simply understood as: When we have a number of options, we will usually select the most beneficial one at sight [47][48], as illustrated in the following. Figure 4.2 shows a simple route selection.

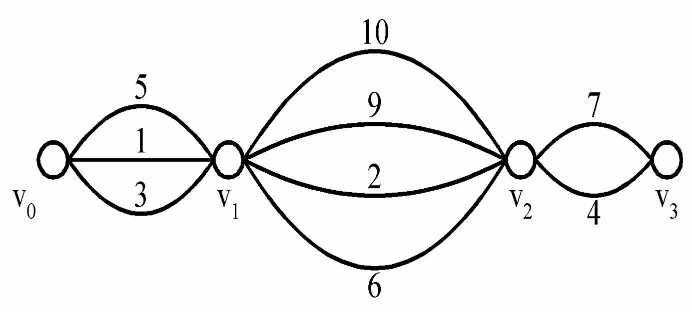

Figure 4.2: A simple route selection

Suppose we are to choose the most popular route from V_0 to V_3, Greedy Algorithm would suggest 5 + 10 + 7 = 22, which is the common one. This can not be an easy job every time. The following figure 4.3 explains why:

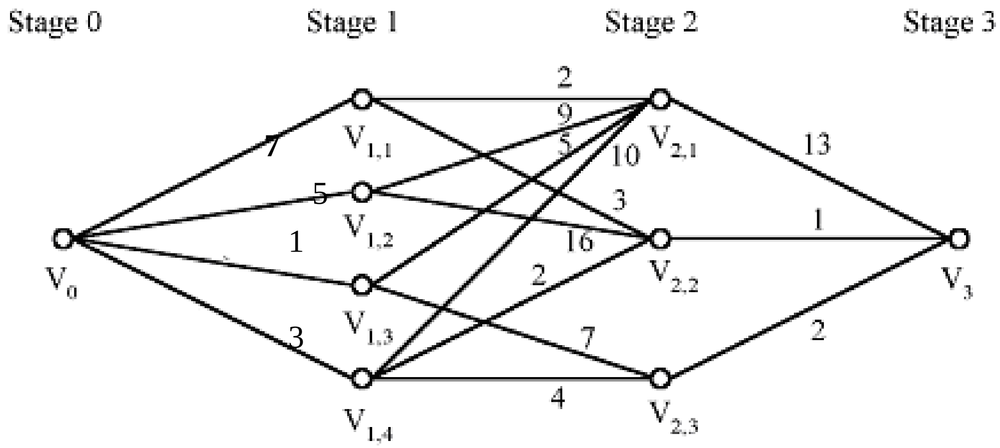

Figure 4.3: A complex route selections

Likewise, we can choose the most common one from V_0 to V_3 and Greedy Algorithm creates a result that indicates the following:

$$V_0 \xrightarrow{7} V_{1.1} \xrightarrow{3} V_{2.2} \xrightarrow{1} V_3 = 11$$

However, 11 is not the most popular route, as the most popular access from $V_0 \rightarrow V_{1.1}$ does not necessarily mean that $V_{1.1}$ will surely lead us to the most popular route. To solve this problem, we may use the following methods.

(1) The first method is to calculate each and every stage of possible routes and to choose the largest one. This could be an easy and straight forward method, but this means heavy load when there are many routes available. In this method, it would take as many as 9 calculations before locating the best route.

(2) With the second method, we have to set up first the Table 4.2 given as follows and arrange the stage routes from large to small by count.

Table 4.2: Greedy Algorithm routing Table

Stage0-Stage1

Session	Count
$V_0 \rightarrow V_{1.1}$	7
$V_0 \rightarrow V_{1.2}$	5
$V_0 \rightarrow V_{1.4}$	3
$V_0 \rightarrow V_{1.3}$	1

Stage 1-Stage2

Session	Count
$V_{1.2} \rightarrow V_{2.2}$	16
$V_{1.4} \rightarrow V_{2.1}$	10
$V_{1.2} \rightarrow V_{2.1}$	9
$V_{1.3} \rightarrow V_{2.3}$	7
$V_{1.3} \rightarrow V_{2.1}$	5
$V_{1.4} \rightarrow V_{2.3}$	4
$V_{1.1} \rightarrow V_{2.2}$	3
$V_{1.4} \rightarrow V_{2.2}$	2
$V_{1.1} \rightarrow V_{2.1}$	2

Stage2-Stage3

Session	Count
$V_{2.1} \rightarrow V_3$	13
$V_{2.3} \rightarrow V_3$	2
$V_{2.2} \rightarrow V_3$	1

Then, by following the order Stage 0 → Stage1 → Stage2 → Stage3, we can look for the corresponding routes and add up to the count values before comparing both of the total count values of each routes at the same level and the maximum potential value of the next level. If both values are large, it indicates that we have located the best route and do not have to continue. In the following example:

Step 1：$V_0 \xrightarrow{7} V_{1.1} \xrightarrow{3} V_{2.2} \xrightarrow{1} V_3 = 11$

Step 2：$V_0 \xrightarrow{7} V_{1.1} \xrightarrow{2} V_{2.1} \xrightarrow{13} V_3 = 22$

Step 3：$V_0 \xrightarrow{5} V_{1.2} \xrightarrow{16} V_{2.2} \xrightarrow{1} V_3 = 22$

Step 4：$V_0 \xrightarrow{5} V_{1.2} \xrightarrow{9} V_{2.1} \xrightarrow{13} V_3 = 27$

When we realize that the fourth total count of 27 is higher than the $V_0 \rightarrow V_{1.2} \rightarrow V_{2.2} \rightarrow V_3=22$ at the same level and the maximum potential value of 3+10+13=26. At the next level, we can be sure that the most popular route is $V_0 \rightarrow V_{1.2} \rightarrow V_{2.1} \rightarrow V_3$ = 27. We have located the best route in the forth step. Now we can save the analysis result in the database for updating the web framework.

4.2 Updating Grid Portal Structure

Finally, we add PHP commands to the web for regular reading of data and linkage in the database, while updating web hyperlinks as a way to adjust the web framework. The following shows a fragment of PHP command:

```
<p><a href="% Firework1_url %"> % Firework1_name % </a></p>
```

Here, Firework1 _url stands for the website of a specific Web file name in the Web server, such as “Firework1.htm”. Firework1_name is a specific record in the database which stores the mining result. Figure 4.4 and Figure 4.5 shows the screenshot of web structure update procedure.

Figure 4.4: Web structure before update

Figure 4.5: Web structure after update

This kind of web updating used to be conducted manually by reading the data first (ranking, for instance) before the administrator's manual updating of hyperlink. From now on, all we have to do is trigger the automated updating system. The web will update its links by following the latest result of mining. Substitution of Web Usage Mining for web ranking will be more of an objective, as Mining result stands for "behavior" of all browsers on the web, not only ranking.

Chapter 5

Conclusion and Future Work

This research expects to apply Grid Computing technologies to integrating the idling computer resources in the schools for saving cost and adequate application for its resources. If it is commonly and widely applied, a huge amount of software and hardware equipment purchasing expenditure can be saved and the goal of complementation and sharing of resources among schools could be achieved. Moreover, the school with insufficient budget can also obtain better services and enormous teaching resources through the Grid technology. With the endless joining up of partner schools, we can also expect the outcome and experiences of this research that can provide reference to the schools, through which they could develop an e-Learning environment for reduction the cost and time in developing a similar system.

To make the resources in the e-Learning grid Platform available to users, we have to set up a Grid Portal. However, as Grid Portal may not be necessarily meet that actual need of different users. We have to use Web Mining for optimization to be able to make a Grid Portal run more efficiently.

Web Usage Mining is a kind of Web Mining that deals with the study of web browsers' routes and the behavior patterns. Since most existent studies tend to focus on improving and unveiling algorithm and efficiency of pattern, we like to propose a method and idea for automated web updating. This will make the framework of Grid Portal adjustable on its own by following the browser's preference. This will not only save significant amount of effort and time in updating webs, but also minimize the errors resulting from subjective judgment. We do hope that in the near future, we could make Grid Portal capable of learning on its own while predicting browsers' behavior and conducting automated adjustment as a Smart-Learning Grid Portal.

Besides the e-Learning platform, there are many Data Grid technologies applications. The following two topics can be further discussed in the future.

5.1 Global Digital Library Grid

Data Grid can be also used to link Digital Collection institutions such as libraries, museums and archives to form an enormous virtual museum [49], which we called Digital Library Grid. Through the internal files management mechanism, Digital Library Grid allows its users to sharing digital resources and be able to conduct a catalog management, replica transfer management and the most efficient access to files. Once scattered resources are linked together, all the end users have to do is to go to a single inlet to search for the files that is scattered all over the world to fulfill its searching requirements. With replica selection, the most efficient site would be located automatically and a copy of the file would then appear in the local site for the end user.

Based on the architecture of Data Grid, information placed in each Local Grid should have its uniqueness and will not be repeated. For example, library A is placed for fine arts and library B is placed for literature. In doing so, it will not only result to its waste resource investment. However, as the connection between libraries will go through the Internet, it is possible that in the course of data transmission, there may be occurrence of abnormality due to unstable connection quality. In order to reduce this kind of condition, we have to separate parts of Data Grid storage spaces as Proxy [50] so that it could to increase the efficiency of data retrieval.

5.2 Access Grid and Data Grid

Access Grid is the technology developed by U.S. Argonne National Laboratory. It enables many people to process interactions and opinion exchanges through video and audio. At present, it has been used in many sites that need video such as training, teaching, conference and seminar, etc [51]. The application software of Access Grid is called Access Grid Toolkit. This is a freeware that anybody can freely download from the Access Grid web site. On the web site, it also provides complete technical archives for giving necessary assistance to its users. Up to September 2003, the latest version of Access Grid Toolkit is Version 2.1 [52].

In Taiwan, the most well-known example of Access Grid application is the use for Anti SARS during 2003 [53]. With the assistance of three researchers from U.S. Argonne National Laboratory, some medical institutes of Taiwan have successfully installed Access Grid in enabling Taiwanese doctors to discuss the SARS patients' condition with the medical specialists of other medical institutes around the world in order to obtain more objective and precise judgment in coping with the spread of SARS virus.

Before the creation of Access Grid Toolkit Version 2.0, Access Grid is only a kind of video conference system. Subsequent to the Version 2.0, Access Grid has already tightly combined with Grid technology. All the functions of Grid technology such as authentication, resource allocation, and remote data access and fault detection are the standard function of Access Grid; therefore the Grid Middleware Globus Toolkit needs to be installed before the installation of the Access Grid Toolkit 2.0.

The audio and video materials that recorded by Access Grid is usually so huge that they require high capacity of storage device such as Disk Array, NAS (Network Attached Storage) or SAN (Storage Area Network). These equipments are very expansive; therefore, we can use Data Grid technology to substitute the above storage device.

References

[1] Stephen Tucker, Andrew Pigou, and Thom D. Zaugg, "e-Learning making it happen now", *Proceedings of The 30th annual ACM SIGUCCS fall conference on User services conference*, pp. 292-293, November 2002**.**

[2] Distance Education and the WWW, http://tecfa.unige.ch/edu-comp/edu-ws94/contrib/peraya.fm.html

[3] Fran Berman, Geoffrey, and Tony Hey, "*Grid computing-Making the Global Infrastructure a Reality*", 2002 John Wiley & Sons Ltd, pp. 9-49 ISBN: 0-470-85319-0.

[4] Alexander Reinefeld and Florian Schintke, "Concepts and Technologies for a Worldwide Grid Infrastructure", *Euro-Par 2002 Parallel Processing, volume 2400 of Lecture Notes in Computer Science*, pp. 62-71, 2002.

[5] What is grid computing? http://www-1.ibm.com/grid/about_grid/what_is.shtml

[6] T O'Hagan C, "The integration of television and the Internet", *Advanced Learning Technologies, 2001. Proceedings IEEE International Conference*, pp. 475 -477, August 2001.

[7] W. Allcock, A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and S. Tuecke, "The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets", *Journal of Network and Computer Applications*, Volume 23, pp. 187–200, 2001.

[8] I. Foster, I. Kesselman, C. Nick. J, and Tuecke. S, "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration", *Globus Project*, 2002.

[9] Introduction to Grids and the Globus Toolkit, http://www.globus.org/training/grids-and-globus-toolkit/

[10] A. Ghiselli, "DataGrid Prototype 1", *TERENA Networking Conference*, June 2002

[11] GriPhyN - Grid Physics Network http://www.griphyn.org/index.php

[12] Particle Physics Data Grid (PPDG) project at SLAC
http://www.slac.stanford.edu/xorg/ngi/ppdg/ppdg-slac.html

[13] The Data Grid Project
http://eu-datagrid.web.cern.ch/eu-datagrid/

[14] Vazhkudai, S. Tuecke, and S. Foster. I, "Replica selection in the Globus Data Grid", *Cluster Computing and the Grid, 2001. Proceedings of The First IEEE/ACM International Symposium on*, pp. 106 -113, May 2001.

[15] Getting started with the Globus Replica Catalog
http://www.globus.org/datagrid/deliverables/replicaGettingStarted.pdf

[16] Susan Malaika, Andrew Eisenberg, and Jim Melton, "Standards for databases on the grid", *ACM SIGMOD Record*, Volume 32, Issue 3, September 2003.

[17] Globus: Grid Security Infrastructure (GSI)
http://www.globus.org/security/

[18] Globus: Replica Management
http://www.globus.org/datagrid/replica-management.htm

[19] Mary Thomas, Maytal Dahan, Kurt Mueller, Steve Mock, Cathie Mills, and Ray Regno. "Application portals: practice and experience", *Concurrency and Computation: Practice and Experience*, Volume 14, pp. 1427-1443, 2002.

[20] ASC Portal Users Guide
https://www.ascportal.org/portal/usersguide.html

[21] Diesel Combustion Collaboratory
http://doecollaboratory.pnl.gov/research/diesel/homepage.html

[22] NPACI HotPage Portal
https://hotpage.npaci.edu/status/hotpage_main_in.html

[23] Marlon E. Pierce, Choonhan Youn, and Geoffrey C. Fox, "The Gateway computational Web portal", *Concurrency and Computation: Practice and Experience*, Volume 14, pp. 1411-1426, 2002.

[24] Uniform Interface to Computing Resources
http://www.fz-juelich.de/unicore/

[25] Jason Novotny, “The Grid Portal Development Kit”, *Concurrency and Computation: Practice and experience,* Volume 14, pp. 1129-1144, 2002.

[26] Sun Grid Engine Software
http://wwws.sun.com/software/gridware/

[27] Thomas, M.; Mock, S. Dahan, M.; Mueller, K. Sutton, and D. Boisseau, J.R. “The GridPort toolkit: a system for building Grid portals”, *Proceedings of The 10th IEEE International High Performance Distributed Computing Symposium on,* pp. 216 -227, August 2001.

[28] NPACI A National Partnership
http://www.npaci.edu/

[29] OGCE: Open Grid Computing Environment
http://www.ogce.org/index.php

[30] D. Gannon, G. Fox, M. Pierce, B. Plale, G. von Laszewski, C. Severance, J. Hardin, J. Alameda, M. Thomas, and J. Boisseau, “Grid Portals: A Scientist's Access Point for Grid Services (DRAFT 1)” GGF Working Draft, Sep. 2003.

[31] Towards Open Grid Services Architecture
http://www-fp.globus.org/ogsa

[32] Gang Xue, Fairman. M.J, and Cox, S.J, “Exploiting Web technologies for Grid architecture, Cluster Computing and the Grid”, *The 2nd IEEE/ACM International Symposium CCGRID2002,* pp. 256 -257, May 2002.

[33] Web Services Architecture, W3C Working Draft 8 August 2003.
http://www.w3.org/TR/ws-arch/

[34] NMI Portal Installation Instructions
http://www.ogce.org/install.php

[35] A. Chervenak, E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunst, M. Ripeanu, B, Schwartzkopf, H, Stockinger, K. Stockinger, and B. Tierney, “Giggle: A Framework for Constructing Sclable Replica Location Services”, *Proceedings of IEEE Supercomputing 2002 (SC2002),* November 2002.

[36] Lizhen Liu, Junjie Chen, and Hantao Song, "The research of Web mining", *Proceedings of the 4th Intelligent Control and Automation World Congress on* Volume 3, pp. 2333 -2337 , June 2002.

[37] Osmar R. Zaiane, Man Xin, and Jiawei Han. "Discovering Web access patterns and trends by applying OLAP and data mining technology on Web logs", *Proceedings of The Advances in Digital Libraries ADL'98 Santa Barbara CA*, pp. 19-29, April 1998.

[38] Ludmila Cherkasova, "Improving WWW Proxies Performance with Greedy Dual Size Frequency Caching Policy", *HP Laboratories Report No. HPL-98-69R1*, CA, USA, Palo Alto, April 1998.

[39] Qiang Yang, Haining Henry Zhang, and Tianyi Li. "Mining web logs for prediction models in WWW caching and prefetching", *Proceedings of The Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD'01*, pp. 473-478, August 2001.

[40] Robert Cooley, Bamshad Mobasher, and Jaideep Srivastava. "Web Mining: Information and Pattern Discovery on the World Wide Web", *Proceedings of The 9th IEEE International Conference on Tool's with Artificial Intelligence*, pp. 558-567, November 1997.

[41] Jaideep Srivastava, Robert Cooley, Mukund Deshpande, and Pang-Ning Tan. "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", *ACM SIGKDD Explorations*, Volume 1, Issue 2, pp. 12-23, 2000.

[42] Ramakrishnan Srikant and Yinghui Yang. "Mining Web Logs to Improve Website Organization", *Proceedings of the 10th International World Wide Web conference* Hong Kong, pp. 430-437, April 2001.

[43] M. Carl Drott. "Using Web server logs to improve site design", *Proceedings of The sixteenth annual International conference on Computer documentation. Quebec Canada*, pp. 43-50, September, 1998.

[44] Myra Spiliopoulou, and Carsten Pohle. "Data Mining for Measuring and Improving the Success of Web Sites", *Data Mining and Knowledge Discovery* Volume 5, pp. 85-114, 2001.

[45] M. Spiliopoulou and L.C. Faulstich. “WUM: A Tool for Web Utilization Analysis”, *Proceedings of The EDBT Workshop WebDB'98, LNCS 1590. Springer Verlag*, pp. 184-203, March 1998.

[46] WUM : A Web Utilization Miner
http://wum.wiwi.hu-berlin.de/

[47] Dictionary of Algorithms and Data Structures
http://www.nist.gov/dads/HTML/greedyalgo.html

[48] Greedy Algorithm -- from Math World
http://mathworld.wolfram.com/GreedyAlgorithm.html

[49] Distance Education and the WWW
http://tecfa.unige.ch/edu-comp/edu-ws94/contrib/peraya.fm.html

[50] Wolfgang Hoschek, Javier Jaen-Martinez, Asad Samar, Heinz Stockinger, and Kurt Stockinger, “Data Management in an International Data Grid Project”, IEEE/ACM International Workshop on Grid Computing Grid'2000, December 2000.

[51] Stevens. R, Papka. M.E, and Disz. T, “Prototyping the workspaces of the future”, *Internet Computing, IEEE , Volume: 7 Issue: 4* , pp. 51-58, July-Aug. 2003.

[52] Access Grid
http://www.accessgrid.org/

[53] Sars Hope
http://www.sarshope.org/