PAREGIN: Pakistan Research and Education Grid Initiative

By

Ovais Ahmad Khan (m9543)

In partial fulfillment of the requirements
for the degree of

Master of Science in Computer Science

Department of Computer Science
National University of Computer and Emerging Science, Karachi Campus

General Information

Name:	Ovais Ahmad Khan
Std #:	M9543
Program:	MS (CS)
Specialization:	Data Communications and Networks
Thesis Advisor:	Dr. Misbahuddin Mirza
Thesis Title:	PAREGIN: Pakistan Research and Education Grid Initiative

Approved by Graduate Committee:

Dr. Misbahuddin Mirza Date
Thesis Advisor (On Behalf of Committee)

PAREGIN: Pakistan Research and Education Grid Initiative

Abstract

Ever since the off the shelf networked computers became popular, the interest of people in distributed systems is increasing. Many architectures, strategies and libraries were made. Grid Computing is one such emerging new paradigm for next-generation distributed computing. It enables the sharing, selection, and aggregation of geographically distributed heterogeneous resources.

A number of teams all over the world are working on various aspects of grid computing. This work is an effort to have a research and education grid in Pakistan considering its specific limitations and requirements. It studies the various implementation architectures of grids and based upon the results, proposes Pakistan Research and Education Grid Initiative (PAREGIN). PARAGIN is collaborative grid architecture suitable for Pakistan and provides guidelines for the middleware, network infrastructure and services useful for researchers, students and teachers. The work also includes cost and usability analysis of the services. Based upon the proposal, a prototype grid was implemented along with a smaller set of services.

Acknowledgments

Research is a long and arduous process, in which I was helped by many individuals. At the outset, I would like to express my appreciation to Professor Dr. Misbahuddin Mirza for his advice during my research endeavor. As my supervisor, he has constantly forced me to remain focused. His observations and comments helped me in choosing the right path during my research work has proved essential to the success of my work.

I would like to thank my former supervisor Professor Dr. Qurban Ali Memon who helped me to start my Masters research and to establish the overall direction of the research and to move forward with in depth investigation.

I would like to thank my other thesis committee members, Dr. Zubair Shaikh and Dr. Asim Ur Rehman for pointing out my mistakes during the thesis proposal phase. Their comments really helped me in completing the research and achieving the results.

I acknowledge the support and love of all my friends and seniors at FAST for making these six years at FAST an enjoyable phase of my life. Finally, this work wouldn't have been possible without the love, support and encouragement from my family members.

The list is by no means complete. There are many people with whom I have engaged in informal discussions which has sub-consciously inspired me to do certain things or made me realize my mistakes.

Table of Contents

List of Figures

Chapter 1. Introduction

Traditionally, most computing systems used to work individually, in a stand-alone manner. Even though they were networked, it is fairly recent that these resources were used in a cooperative manner. The main reason for this was the slower network speeds. With the passage of time, network speeds increased and so did the storage capacities. This along with the availability of off-the-shelf, high-end networked computing resources increased the interest of many in making use of these resources in a coordinated manner. This merging of the individual systems to build a larger cooperative system is termed as a distributed system. A distributed system is thus one in which hardware and software components of networked computers communicate and coordinate their activities by sharing resources such as information, data, compute cycles, bandwidth and storage.

This revolution, coupled with the spread of Internet and the advent of the standardized protocols. Generally most of the resources available via Internet are hosted on servers and accessed using certain protocols like HTTP, FTP, etc. Most compute intensive jobs require supercomputers which have special procedures to submit jobs. The recent trend of building distributed systems gave rise to Grid Computing [1]. Thus it was possible to make use of the otherwise idle resources to solve large scale scientific and industrial problems that require lots of processing power, data storage, and specialized equipments for experiments and analysis. Such advanced distributed systems can also be used for collaboration and sharing of various resources between organizations.

Grid is a distributed system infrastructure for putting together computational data and other resources which are geographically and organizationally dispersed. Grid in this way provide infrastructure for supporting multi-institutional collaboration and resource sharing. The heterogeneity, flexibility and reliability set grids apart from supercomputers, server farms, clusters and peer-to-peer schemes. The other properties of grid includes distributed control, dynamic configuration, adaptability, and quality of service guarantees (see Section 2.2.)

The overall motivation for current large-scale, multi-institutional grid projects was to enable resource sharing and facilitate human interactions. Currently, various grid projects are being developed for research and commercial purposes. These include large-scale science and engineering projects such as Grid Physics Network (GriPhyn), Access Grid, NASA Information Power Grid, EU Data Grid, TeraGrid and many others (see Appendix I). Many grid middleware and supporting tools like resource managers and brokers are also being developed like Globus, Gridbus, and others. With so much work going on in this direction, the need for a standard body arose which can discuss and make various standards and protocols. Thus the Global Grid Forum (GGF) was formed by various stake holders including large enterprises, research organizations and universities. GGF has already made various standards and many others are being drafted by its working groups.

This report describes the research carried out as part of MS thesis to investigate Distributed Systems. The aim of the study was to identify and work on the business needs, architectures and approaches needed for developing a distributed system for academic collaboration in a developing country like Pakistan. Due importance in this study was given to the computing and networking infrastructure and the needs of the local universities and research organizations. The following objectives were set at the start of this thesis study:

- **Implementation of an experimental grid:** A Campus-Grid will be established in the campus of National University of Computers and Emerging Sciences in Karachi.

- **Grid Architectures:** Study of various different architectures would be conducted along with their respective implementation scenarios and proper model for implementation in local setting would be proposed.
- **Policies for the grid:** All organizations which operate or utilize grids do so for some business purpose. We need a policy document so that the grid environment supplements the business goals of the company. Result of the policy will be strict control over the use of resources and monitoring the behavior of the individuals. An effort will be made to address this issue in the main document.
- **Wireless Grids:** This study will also review how wireless interfaces to grid services. Some interface specific issues will be discussed and approach and architecture for such an access will be proposed.

The rest of this thesis is organized as follows. Chapter 2 focuses on the basics of the grid and introduces the key concepts related to the grid. It also presents the various implementations of grid in collaborative and related applications along with their architectures. Chapter 3 discusses the proposed grid implementation architecture suitable for a developing country like Pakistan, while chapter 4 touches upon the implementation issues. Chapter 5 contains a discussion related to the various problems that hinder the widespread acceptance of grid. Chapter 6 presents conclusion and provides direction for future work in this direction.

Additional information is provided in Appendices. Appendix I contains information regarding a few of the grid related projects. Appendix II shows the current status report of PERN network as shown on its website.

Let us now look at the current state of affairs in distributed systems specifically the grid and currently available networking infrastructure in Pakistan.

Chapter 2. Distributed Systems

In this chapter, we present the background of distributed systems especially the grid along with a discussion on existing and proposed network infrastructure in Pakistan.

2.1 Introduction

A Distributed System can be defined as one in which hardware and software components at networked computers communicate and coordinate their activity by sharing resources such as information, data, compute cycles, bandwidth and storage. Examples include the World Wide Web, Clusters, Mobile Computing, Grid Computing, and many others.

Historically, workstations and personal computing systems used to work individually solving smaller problems while supercomputers were used for complex problems. With the increase in the network speeds and storage density, coupled with the availability of high-end computers and communication equipments at reasonable prices, this trend has changed during the last few decades. It is now agreed upon that we are living in an exponential world with transistor count doubling every 18 months (Moore's Law), storage density doubling every 12 months, and fiber optics speeds doubling every 9 months (see Figure 1). This shows that Network speeds will grow at double the rate of processor speed, thus the case for distributed computing becomes even stronger.

Work in similar direction can be found dating back to as early as 1969. In 1969, Leonard Kleinrock, was quoted as saying *computers are still in their infancy, but as they grow up and become more sophisticated, we will probably see the spread of 'computer utilities', which, like present electric and telephone utilities, will service individual home and offices* [2].

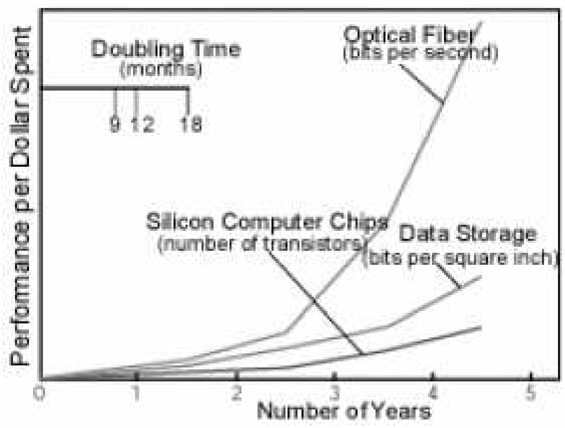

Figure 1: Moore's Law vs. storage improvements vs. optical improvements. Graph from Scientific American (Jan-2001) by Cleo Vilett.

MetaComputing [3] was used in early 90's to describe a similar concept by Smarr and Catlett. They defined MetaComputing as *the computing resources transparently available to the user* [3]. In the beginning Ian Foster also used this term to describe Networked Virtual Supercomputer which uses high-speed network to connect various devices [4].

The term Grid was first introduced by Foster and Kessler in 1998 to define computational grid as *hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities* [5]. They later generalized the definition to all resources, terming it as *coordinated resource sharing among dynamic individuals, organizations and institutions* [6]. Grid computing is one of the most important fields which emerged during the last few years as a result of efforts of various people. With so many contributions from Ian Foster ranging from naming this concept as the Grid to providing architecture and implementation for the grid, he is known as the Father of Grid Computing.

Since this work is mainly focused on grid like architectures from the distributed systems family, so we will focus on Grid Computing from now onwards.

2.2 The Grid

Grid can be defined as *an infrastructure that supports large scale, coordinated, resource sharing among heterogeneous systems that span institutional or organizational boundaries in a dynamic manner* [7].

It is to be noted that there exist a lot of controversy between what is grid and what is not. But essentially the grid can be characterized by the following three point checklist provided by Foster [8]:

- Coordinated resources not subject to centralized control
- Uses standard, open, general purpose protocols and interfaces
- Delivers non-trivial Quality of Service

Let us look at some of the other aspects of Grid Computing. First we will look at the major properties of a Grid System [9]:

- **Heterogeneity:** There can be any type and number of resources in a Grid. The resources may stretch across administrative boundaries. These resources may include a number of processor, memory and network architectures.
- **Scalability:** The grid and the applications built on top of it need to be scalable because of the possibility that many times more resources can be added to the Grid as it evolves.
- **Dynamic Configuration:** The machines and resources can arrive and leave, thus enabling the administrators to make use of their resources only locally when needed, while at the same time sharing it when they are idle.
- **Adaptability:** In the presence of so many resources, failure would be quite common. Applications and resource managers must be able to handle this.

- **Distributed control:** No central control of all resources, neither the resources are available in a peer-to-peer manner. All organizations can have their own policy of sharing the resources.
- **Quality-of-service maintenance:** The Grid system must deliver non-trivial Quality of Service. Workloads are balanced, with attention to task priority, not just best-effort rules.
- **Bandwidth Diversity:** With so many different organizations interconnected, there exists diversity in the bandwidth available at various places.

2.3 Uses of the Grid

Grid was primarily devised for computationally intensive scientific work like systems design, data analysis and visualization, climate research, and so on (see **Error! Reference source not found.**). Hence it was used for these purposes in various projects. Later on the concept was generalized to sharing just about any kind of resource. The new capabilities provided by the grid greatly facilitate the scientists and engineers. Some of the ways grid can be used are [5, 11]:

- High throughput computing
- Distributed Supercomputing
- Data-intensive computing
- On demand computing (utility computing)
- Collaboration within and among organizations
- Data sharing between various offices and partners
- Virtual enterprises and virtual markets could be developed easily using grids
- Providing virtual services, like virtual lectures, virtual stores, etc
- To connect and share heterogeneous distributed resources

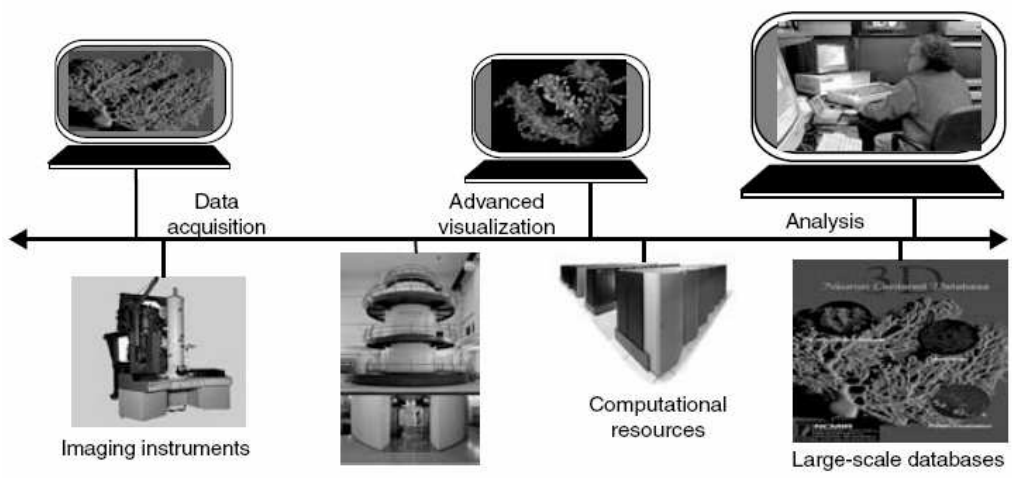

Figure 2: Grid resources linked together for a Telescience application (http://www.npaci.edu/Alpha/telescience.html) [10]

There are many real world grid uses and implementation and these include GriPhyn, Access Grid, TeraGrid, NASA Information Power Grid, EU Data Grid, and many others. More detailed discussion regarding these projects is included in Appendix I.

2.4 Classification of the Grid

Based on the way grid is used, there can be several types of grids:

- **Computational Grids:** These grids provide secure access to huge pool of processing power suitable for high throughput applications and computation intensive computing.
- **Data Grids:** Data grids provide an infrastructure to support data storage, data discovery, data handling, data publication, and data manipulation of large volumes of data actually stored in various heterogeneous databases and file systems.
- **Utility Grids:** Just like data and computation, software or just about any resource can be shared using grid. The main services provided through utility grids are

software and special equipments. For instance, the applications can be run on one machine and all the users can send their data to be processed to that machine and receive the result back.

- **Collaboration Grids:** With the advent of Internet, there has been an increased demand for better collaboration. Such advanced collaboration is possible using the grid. For instance, persons from different companies in a virtual enterprise can work on different components of a CAD project without even disclosing their proprietary technologies.

Another classification is given in [7] namely, resource sharing tools and organizational tools. But it seems difficult to agree to this division as the enterprise, partner and utility Grids mentioned in the paper as organizational tools, are evolutionary stages towards utility computing. Also if we consider these grids as such then we also have to take into account the evolutionary architecture of cluster, inter and intra grid.

2.5 Relationship with other technologies

As evident from previous description, grid has strong relationship with the older technologies especially those related to distributed systems and service oriented architectures. The following are some related technologies along with their qualities and comparison with the grid.

2.5.1 CORBA

Common Object Request Broker Architecture (CORBA) is an open distributed object computing standard developed by the Object Management Group (OMG) [12]. CORBA automates most of the common programming chores associated with development of distributed applications. Even though CORBA provides a rich set of services, its initial design focus on intranet applications is making it difficult for inter-organizational deployments.

2.5.2 Jini & RMI

Jini Technology [13] is architecture for the construction of systems from objects and networks. Jini is primarily concerned with communications between devices (not what devices do). It provides simple mechanisms which enable devices to plug together to form a distributed computing environment. The system so formed is called Jini Community in which the devices are tied together using Java Remote Method Invocation (RMI).

2.5.3 Peer to peer computing

Peer-to-Peer (P2P) computing [14] is an emerging computing model in which machines share resources and coordinate without any central servers. Many applications are using this model like Freenet, Napster, Grokster, eDonkey, and many others.

A detailed discussion on comparison and integration of P2P and grid systems is contained in [15]. Basically, it explored on how a grid can be implemented in a P2P model instead of the traditional layered model of the Web.

A very detailed relationship between grid and some other technologies like WWW, ASPs, etc is given in [6].

2.6 Setting up a Grid

Setting up a grid involves a lot of different steps. Even the simplest of all grids would require a very good network infrastructure, storage and computational resources as well as the middleware and resource managers. The other features can be added as required and desired.

2.6.1 Infrastructure Requirements: The Network

Network is the heart of any successful grid implementation. Most grid implementations are using the high speed fiber optics backbone providing speeds up to 40 Gbps. If the network connectivity is not reliable or appropriate bandwidth is not available, the grid applications deployed on top of it should communicate as little as possible and support high latency, or the deployment will fail. For normal data-intensive or low latency application, high speed fiber optics network should be there.

At present, grids are built on high-performance network. According to [10], by the year 2002, most of those networks had roughly 10Gbps backbone. Examples of such networks include Abilene Network (USA), SuperJanet backbone (UK), GEANT Network (intra-Europe), APAN (Asia Pacific), and others. A given institution is connected via about 1 Gbps link to the backbone and has around a 100 Mbps LAN. This means that there is a 10:1:0.1 Gbps ratio between national, organizational, desktop links.

A Global Terabit Research Network (GTRN) [22] is expected to enhance this ratio many times. GTRN aims to enhance the international, national, organizational, optical desktop and copper desktop link ratios to 1000:1000:100:10:1 Gbps, a 100 times increase from the current state. The NSF funded TeraGrid project [16], has one of the most striking network infrastructure with four locations spread across USA linked with a backbone of 40 Gbps.

2.6.2 Hardware Requirements: Shared Resources

Resources are the main building block of the grid. The whole concept was introduced to share resources, so there should be redundant and highly-available resources. Even simple grid systems include a combination of high speed and high capacity data storage in addition to a large amount of computational power. Additionally, a scientific grid consists of specialized analysis, visualization and scientific equipments. Other specialized grids have other components as well.

2.6.3 Software Requirements: The Grid Middleware

The software layer between the operating system and the applications is termed as middleware. It provides a variety of services required by an application to function correctly. Middleware has recently re-emerged as a means of integrating software applications running in distributed heterogeneous environments. Middleware thus refers to the software which is common to multiple applications and builds on the network transport services to enable ready development of new applications and network services. CORBA, for example, defines a middleware standard.

In a grid, the middleware is used to hide the heterogeneous nature and provide users and applications with a homogeneous and seamless environment by providing a set of standardized interfaces to a variety of services [9]. With the use of service oriented architecture in grid computing (see next section for details), this middleware consists of the services commonly used by the grid applications like authentication, resource access and management, etc. Examples of such middleware include the Globus toolkit.

With so many resources available, these need to be looked up by probable users and need to be managed properly. This task is performed by resource managers which manage resources like processing power by distributing it among the many applications depending upon their priority. GRAM, the Globus Resource Allocation Manager is one such resource manager which is an integral part of the Globus toolkit.

Most of the grids would also require specialized software applications making use of the available grid resources in the most optimal way. These applications are usually built using the services provided by the middleware, resource managers, job schedulers and other components.

2.6.4 Grid Portals: The User Interface for the Grid

In order to provide easy access to the grid services and the resources, a web portal like interface to grid was introduced called The Grid Portal. Just like a web portal allows users to access various resources via a web interface, a grid portal provides access to grid resources. Grid portal utilizes the web browser as a thin client and thus has the advantage of having minimal setup time on the machines of the users. A typical grid portal provides functionality to authenticate users, permit them to access remote resources, help them make decisions about scheduling jobs, and allow users to access and manipulate grid enabled databases and file systems. Grid portal access can also be personalized by the use of profiles, which are created and stored for each portal user.

Many teams have come up with various applications and projects helpful in the development of the grid portals which include NPACI Hotpage, OGCE, SDSC Grid Port Toolkit, Mississippi Computational Web Portal, Lattice Portal, Grid Portal Development Kit and many others.

2.6.5 Grid Certification Authority

The Grid Security Model as described in Globus Grid Security Infrastructure (GSI) is an extension of public key infrastructure (X.509 certificates). In the short term the user generates his short-term proxy using his long-term certificate. Grid implementations thus require presence of a Certification Authority to issue certificates to users and hosts. Grid Certificates are just like the normal certificates used on the Internet and even the same authorities can be used. But due to security considerations, it is expected that grid has its own certification authorities. Most countries working in this area already have such authorities in place.

2.7 The Pakistan Education and Research Network

With network being the heart of the grid, a network infrastructure is the first step in order to set up a Grid Environment. Since this study is regarding grid implementation in Pakistan, we have to look at the Pakistan Education and Research Network (PERN), a nationwide educational intranet connecting premiere educational and research institutions of the country [17]. PERN is designed for collaborative research, knowledge sharing, resource sharing, and distance learning by connecting people through the use of Intranet and Internet resources. Though late, this is a good initiative for Pakistani universities and research institutions.

Even though the networking requirement of most real grid applications are much more than provided by PERN, a developing country like Pakistan can not afford to have another dedicated high-speed network. So, PERN would be our starting point to any real grid computing implementation in Pakistan. The next chapter discusses this implementation and proposed changes in more details.

2.7.1 PERN Architecture

PERN utilizes the existing Optical Fiber System (OFS) of Pakistan Telecommunication Private Limited (PTCL) and National Telecommunication Limited (NTC) and IP/ATM backbone of NTC is utilized for the core network of PERN. The network design of PERN consists of three nodal points at Islamabad, Lahore and Karachi as shown in Figure 3. Most educational institutions are connected to their respective nodal point by a 256 Kb/s to 6Mbps link from the nearest exchange of NTC/PTCL using OFS, DXX (or better system), DRS or VSAT, whichever is technically feasible.

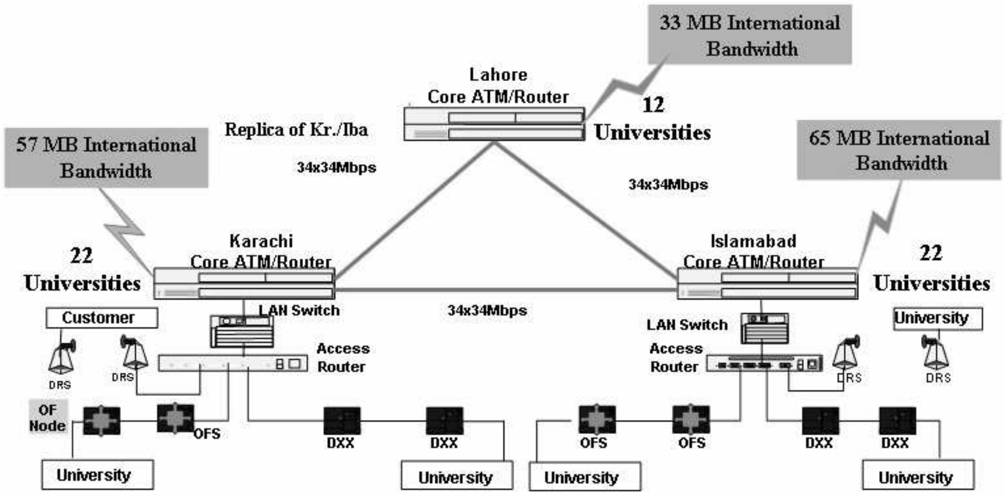

Figure 3: PERN Architecture [18]

According to PERN website, the current interconnection of three main nodes is on 34x34 Mbps and Internet connectivity is 65 Mbps for Islamabad, 33 Mbps for Lahore and 57 Mbps for Karachi [18]. This architecture allows institutions to pool resources with each other through national fiber network and to access Internet from the respective nodal points. The bandwidths provisioned and planned through the PERN network for various universities as mentioned on PERN site is shown in Appendix II.

In order to elaborate the differences between PERN and similar networks of developed countries, the following is the architecture of the United Kingdom National Backbone Research and Education Network (see Figure 4). The figure especially shows differences in speed and reliability of the two networks.

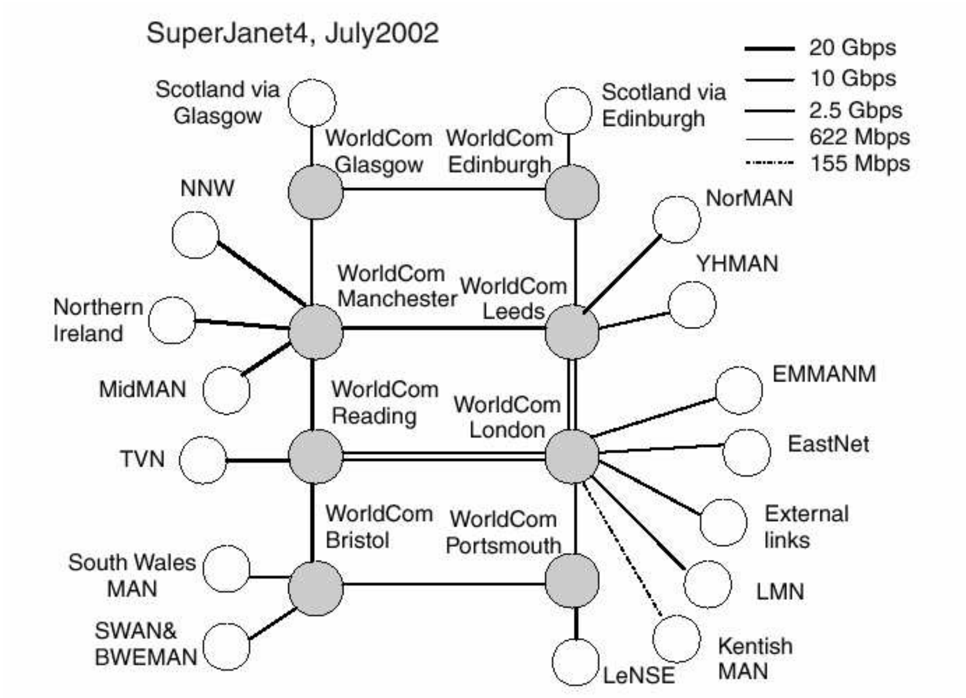

Figure 4: United Kingdom National Backbone Research and Education Network

2.8 Basics of a Grid Architecture

This section includes a discussion on various grid architectures especially the way grid middleware and applications have been implemented. Let's have a look at the basics of Grid Architectures. Grid architecture is basically a layered, service-oriented one with all resources and functions viewed as services. Let us now look at the layered grid model and the service oriented architecture.

2.8.1 Layered Grid Model

As described earlier the grid architecture is layered. The grid community is gradually converging on a layered model which called the Community Grid Model [10]. It provides abstraction of the grid. Figure 5 illustrates this layered model.

There are four horizontal layers each providing abstraction to the layers below them. The bottom layer consists of the resources to be managed and provided by the grid like computers, data storage devices, scientific equipments, etc. The second layer consists of services and software that virtualizes the resource access. OGSA standard and Globus software are fast becoming the de facto standards for this layer providing heterogeneous dynamic resources. The next horizontal layer contains the services commonly used by the grid applications like authentication, resource access and management, etc. The top layer is the actual grid application layer. The vertical layers show the next steps in the development of the grid like addition of wireless devices, sensors, policy and economic models and so on.

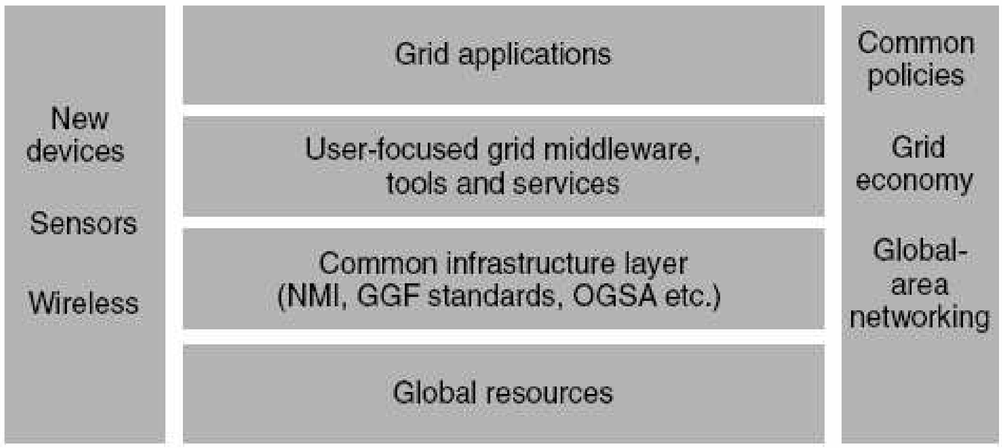

Figure 5: Layered architecture of the Community Grid Model

Figure 6 shows the mapping between the networking protocols and the grid protocols. The difference between the community grid model and this one is that the middle ware is further divided into resource and collective layers.

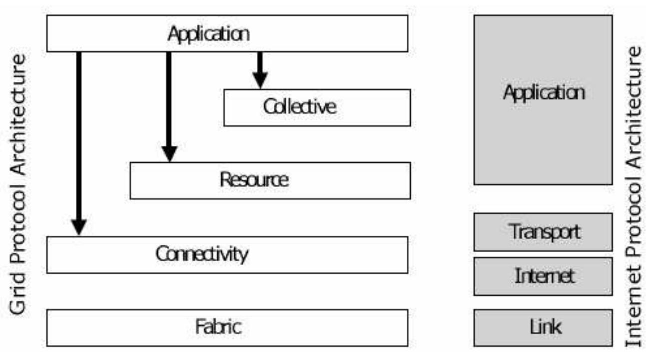

Figure 6: The layered grid architecture and its relationship to the Internet protocol architecture.

2.8.2 Service Oriented Architectures

A service can be defined as a unit of work (some business process) done by some provider. They are loosely coupled, have well-defined, platform-independent interfaces, and are reusable. Since services define only the sequence of message exchanges in order to perform the operation, the actual implementation is flexible. A service Oriented Architecture (SOA) is one in which all entities are services and thus all operations are the result of message exchanges [19].

The most basic SOA implementations include a provider and consumer, while most useful ones also include a service registry (See Figure 7). Web Services and the OGSA (Grid Services) are the related instances of SOA.

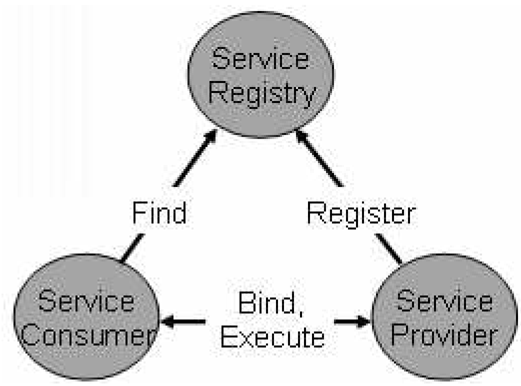

Figure 7: Service Oriented Architecture Actors

2.8.3 The Open Grid Service Architecture

Like any budding technology, grid computing is also facing a shortage of protocols and standards, which are necessary for interoperability. This gap was filled by The Open Grid Services Architecture (OGSA) [20], which is now a Global Grid Forum (GGF) standard. OGSA focuses on the services and thus we may call it a SOA and it derives heavily from the Web Services (WS) and Simple Object Access Protocol (SOAP). OGSA virtualizes the resources, to shield users from their complexity and eventually allow on-demand delivery of computing services.

The objectives of OGSA include resource management across various platforms, delivery of seamless QoS, autonomic management, defining open interfaces and standards using the existing ones [21]. The OGSA main architecture is shown in Figure 8 in which each layer corresponds with the ones shown in Figure 5, namely resources, middleware supporting web services plus the OGSI extensions, OGSA services and the grid applications exploiting the layers below.

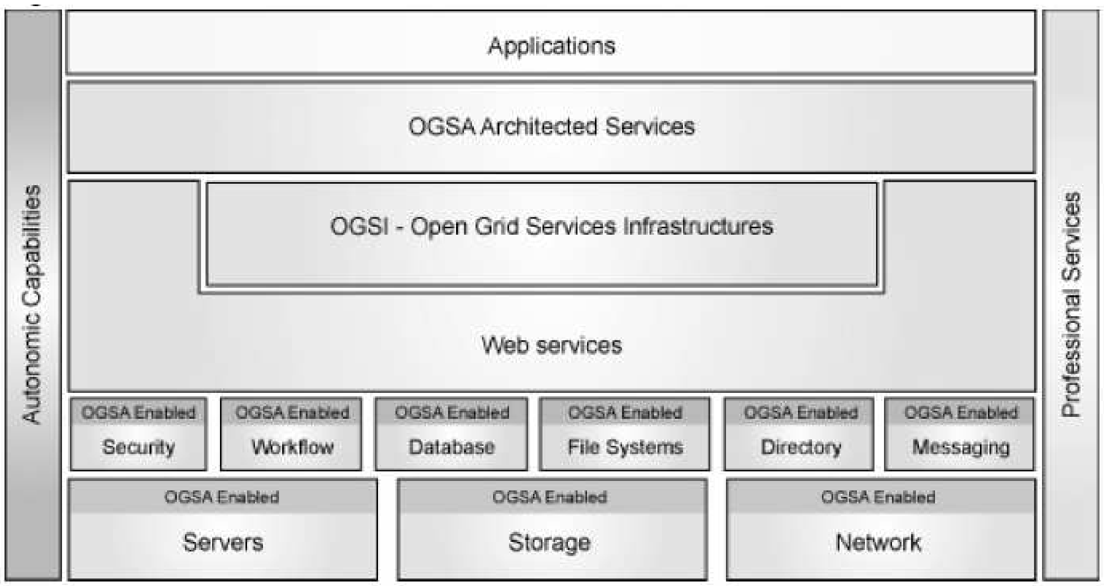

Figure 8: OGSA Main Architecture [21]

OGSI defined a Grid Service [22], which is a Web Service conforming to specific interfaces and conventions. It is designed to operate in a grid environment, and meets the requirements of the grid(s) in which it participates. The deviations from the Web Services are due to the transient nature of the grid services.

OGSA like the most of the layered architectures follow the principle of the hourglass model as shown in Figure 9. The narrow neck of the hourglass defines a small set of core abstractions and protocols like the resources and connectivity protocols in OGSA. The top of the hourglass contains the high level behaviors, while the bottom contains the underlying technologies or the physical resources in OGSA.

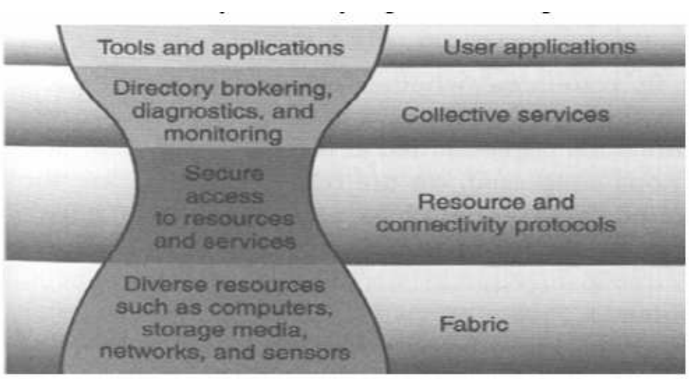

Figure 9: The hourglass Model applied to OGSA

From the above description of OGSA, we can see that the main OGSA components are:

- Open Grid Service Infrastructure (OGSI)
- OGSA Services - Web Service adhering to OGSI
- OGSA Schemas – derived from WSDL

The reference implementation for OGSA standard is the Globus Toolkit 3 (GT3), while many others are trying to conform to the OGSA. Globus Toolkit is a software toolkit available under open source license that addresses key technical problems in the development of grid enabled tools, services, and applications. It enable incremental development of grid enabled tools and applications by offering a modular set of features.

2.9 Evolutionary Architectures of the Grid

2.9.1 Towards a Global Grid

W. Gentzsch advocates an evolutionary architecture for organizations to implement grid computing. He named the three overlapping stages as Cluster, Campus and Global Grids respectively [23].

Cluster Grids: These are local clusters deployed on departmental basis and owned by a single entity.

Campus Grids: Various cluster grids are merged to form a campus / enterprise grid. They can thus have multiple owners.

Global Grids: A number of organizations can decide to merge their campus grids to form a global grid.

This same evolutionary architecture has been named as 3-tiered architecture in [7] which calls these three as Cluster, Intra and Inter Grid respectively. This architecture has been translated to support wireless grids [24]. The basic idea is that implementation of grid is dissolved into various scales, for instance departmental, campus-wide, enterprise, and internet. Different people will use different terminologies, but the core concept will remain the same.

2.9.2 Towards Utility Grids

Similar stage-wise implementation of utility grid has been proposed by Platform Computing. They estimate that the adoption of grid in enterprises would take place in a three-phased evolutionary process [25]. They have named the three phases as Enterprise, Partner, and Utility Grids respectively. In the first phase, enterprises will develop grid within the organization. According to another platform computing article, this process is already continuing since late 90s with large enterprises like GM, Toshiba and others being the first few [26]. Partner Grids is the second phase in which partner companies will link their enterprise grids, using it for collaboration and sharing various resources via it. This is a cut down version of the Global Grid discussed in the previous section, in which enterprise grids of only a few enterprises are linked. The last phase which is still 3 to 5 years from now [26] is the utility computing. In this last phase computing resources will be provided by the service providers over the Internet.

2.10 The Future Businesses

There are many ways that businesses will be transformed once the grids are common-place. Grid Services, Virtual Enterprises, Virtual Markets are some of the future form of the businesses.

2.10.1 Grid Services and Utility Computing

Once we have mature grid technologies, it will not be a far fetch idea to provide IT services just like we use power or other utilities. It is different from grid in the sense that grid is just the infrastructure while the utility computing is a business concept.

Grid Service is a Web service that provides a set of well-defined interfaces and that follows specific conventions. The interfaces address discovery, dynamic service creation, lifetime management, notification, and manageability; the conventions address naming and upgradeability [20]. The current open standard is the Open Grid Services Architecture (OGSA). Standards are further discussed in the next section.

Even though, there is much hype about utility computing, there is no standard definition for the term. It means different things to different people. This is to the extent that the same concept is called on-demand computing, grid services, service-based computing, etc. There is no consensus about the actual definition and the term to be used. Organizations would be able to play various roles in the evolving utility computing market. Some of these roles as identified by [7] are:

- **Grid Resource Suppliers:** These are the companies which supplies grid resources so as to enable Utility Service Providers.
- **Grid Infrastructure Suppliers:** Infrastructure for grid includes hardware infrastructure and middleware.

- **Utility Service Providers (USPs):** They are IT outsourcers in a new name. They would get the required resources and infrastructure from the suppliers and would provide IT resources as services to the users. The major difference from outsourcing is the mode of payment is which now based on usage and adherence to the terms of the Service Level Agreements.
- **Re-Sellers:** They can be optional players in this market. They get the service from USPs and sell it to the users after appropriate repackaging.
- **End Users:** Business users form the bulk of the user-base both in terms of scale and usage. The other user group is leisure users which has sporadic and comparatively low requirements.

2.10.2 Virtual Enterprise

Virtual Organization (VO) or Virtual Enterprise (VE) is not a recent concept. It is just that using grid it seems like it will be accepted much quickly. There are various definitions of Virtual Enterprise but the one proposed by Venkatraman and Henderson as quoted in [28] is rated the best: *A Virtual Enterprise is a network of several companies, which contribute their core competences and share resources such as information, knowledge and even market access, in order to exploit fast-changing market opportunities. The relationship can be long or short term.*

The main idea behind VO is that organizations concentrate on the field in which they are strong leaving other fields for their business partners, thus forming a Virtual Organization. Some people do not consider the Virtual Enterprises as the same as Virtual Organizations because the initial drivers for Virtual Enterprise were simple document sharing. VOs can come into existence for a variety of purposes and vary tremendously from each other. Some differences as listed in [6] are purpose, scope, size, quality, structure, community, and sociology.

The sharing of resources need to be highly controlled with good authentication and authorization mechanisms. Existing technologies for distributed computing and collaboration either lack this control or lack the diversity of resources that can be shared using grid. Another quality of such a sharing is that it is completely ad-hoc in nature. This will eventually lead to on-demand sharing and access of resources and enable organizations to shop around for the best available provide of that resource.

Chapter 3. Proposed Architecture of PAREGIN

This chapter presents the architecture of Pakistan Research and Education Grid Initiative (PAREGIN), a nationwide academic grid for Pakistan. The architecture is basically targeted at educational institutions and research organizations with the main aim being increased collaboration in research and teaching activities as well as sharing expensive resources.

As Pakistan has limited resources and infrastructure especially in educational and research fields, PAREGIN makes optimal use of existing resources and facilities.

This chapter also looks at various issues which were encountered in such an implementation.

3.1 Aims and Objectives

The main beneficiaries of PAREGIN are the researchers and the university administration. Due to the increasingly complex computation and large amount of data, researchers are always short of computing capabilities, data storage devices, as well as expensive equipments for experiments and analysis. Administratively speaking, this would lower the cost of equipments purchased as well as reducing workloads of IT staff.

Some benefits of grid computing for Universities are mentioned in [29]. Derived from there, the following lists various ways in which our universities can benefit from such a grid infrastructure.

- Access to shared computing, data sources, networks and instruments such as sensors and radars to achieve more efficient research as well as making efficient use of under-utilized resources, e.g. all of our campuses have separate digital library subscriptions / e-book collection, separate libraries, separate software bank and separate repository of important documents. If a grid infrastructure is in place, most of these resources could be utilized collectively by all students.

- Collaborate with students and faculty members working on same projects at other campuses in even more effective way - sharing data, documents, results, etc
- Provides a dynamic learning environment for students
- Allows sharing of teaching material and thus provide a more effective, widely available learning setting that enables faculty members to extend their reach
- Mechanisms that allow staff and administrators to offer a broader, better range of services while meeting the demands of constantly changing workloads
- Communicate effectively and easily using integrated video, voice and text messaging coupled with analysis and experimental data shared by each participant. This is of great importance to geographically distributed teams working on same of similar projects.

A multi-campus university can additionally benefit by sharing computing resources more aggressively and thus provide each student and faculty member with resources at all campuses, with little duplication of services and materials (all kind of resources – data, CPU cycles, etc). This coupled with the resilient nature of the grid results in highly-available network with backup and data-storage facilities distributed and replicated at multiple locations.

3.2 The Proposed National Grid

Based on the work of [24] and [7], a proposal to deploy a National Grid consisting of the grids of various educational institutions and organizations of Pakistan is described in this section. This approach is independent of the physical medium used, but will need reliable high-speed network infrastructure. This approach is also extendable to the wireless ad-hoc networks.

In the first step, various campuses, departments or schools of the participating institutions would simply deploy Local Clusters of their own. All these Cluster Grids would be under the administrative control of the respective departmental administrators. As an example we will consider our University (NUCES-FAST) as an institution wishing to develop a University Grid. The University has campuses in four cities with two campuses in Karachi. As a start, each of the five campuses could implement a grid of their own (See Figure 10).

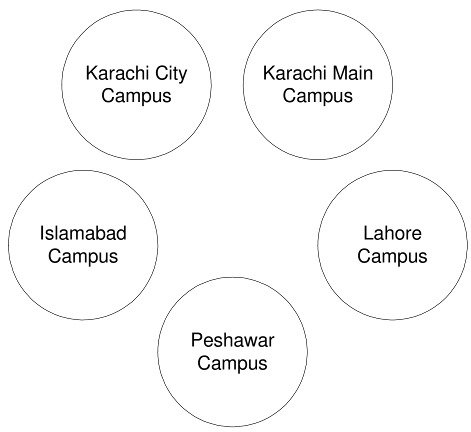

Figure 10: Towards a National Grid - Step 1: Local Clusters deployed on a departmental / School / Campus wide basis

As a second step, we could link the Cluster Grids of the same institutions to form an Enterprise or University Grid. These campuses already have sufficient internet connectivity free during off-peak hours. We could use VPN to connect the grids of these campuses or just connect the machines via the Internet. This would result in a University Grid also known as Campus Grid (See Figure 11). The encircling circle depicts a combined university grid which is also known as Intra-Grid to show that it is internal to a university or institution. This grid would have distributed control but would mainly serve the same university or organization. In our example, the other universities and organizations will develop grids based on their own requirements and resources.

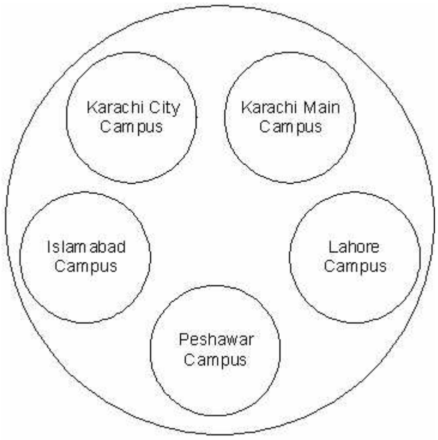

Figure 11: Towards a National Grid - Step 2: The Cluster Grids are merged into Inter-Grid that is within a university or enterprise

The last stage would be to combine these grids and form a National Grid called Inter-Grid to depict the involvement of multiple organizations; even though this is not on a Global Scope but on a National Scope. This Scenario is depicted in Figure 12.

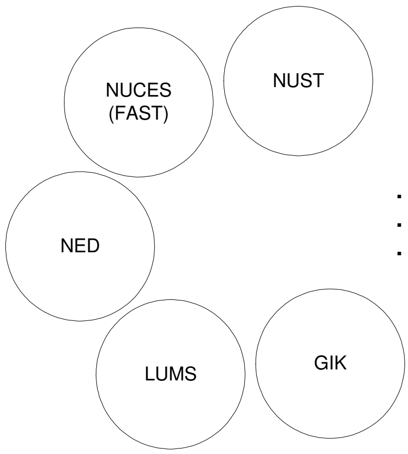

Figure 12: Towards a National Grid - Step 3: Merging the University / Enterprise Campus-Wide Grids of various universities and organizations into a National Grid

3.3 The Architecture

As discussed in section 2.6.1, Network infrastructure plays a vital role in any successful grid implementation. It is to be noted that some grid applications require a dedicated high speed network free from other kind of traffic, while some others can be easily run over a shared network. Thus, networking requirements are heavily dependent on the type of services to be run on a grid.

We can thus make use of Internet / VPN as the network infrastructure. That is very good choice if sufficient Internet connectivity is already available. Thus for a sample grid implementation, we can either use the Internet or a VPN, if available. Using the Internet might be cheaper for starting a grid implementation but is not feasible in a developing country where limited Internet bandwidth is present in most institutions.

A similar solution is to make use of local high speed network. But other types of traffic are present and may cause some quality of service issues for the grid service in case of congestion. Such networks are common in developed countries, but in Pakistan no such real high speed network exists. The best initiative in this direction is PERN (See previous chapter for more details). As part of PERN, a 4Mbps link is proposed between the institutions and the backbone while the main backbone is just above 1 Gbps shared with other NTC users.

Link	Average Speed
National	1 Gbps
Metropolitan	2 Mbps
Organizational	1 Gbps
Desktop	100 Mbps

Table 1: Average link speed of PERN Network

In PERN, we have higher desktop links than the backbone and other links. Such a ratio would not benefit the grid related activities at all as the bandwidth is limited at the higher level. The main bottlenecks are the MAN to Campus links as shown in Table 1. The other problem is the shared nature of this network since this is using the NTC core backbone which is also used by some government organizations and telephony traffic.

If we intend towards the ideal scenario, the ratios should be approaching the values shown in Table 2. This conclusion is based on the discussion of previous chapter, analysis of networking needs of the proposed services, cost analysis, and a survey of the production grids.

Link	Average Speed
National	10 Gbps
Metropolitan	1 Gbps
Organizational	1 Gbps
Desktop	100 Mbps

Table 2: Ideal network needs of PAREGIN

Another approach used by some well-funded grid projects is to make use of a dedicated high-speed network for the grid, similar to the ones described in the last chapter. This has the best network performance amongst all the available choices as only grid related traffic is there.

Even though the networking requirement of most real grid applications are much more than provided by PERN, a developing country like Pakistan can not afford to have another dedicated high-speed network. So, PERN would be our starting point to any real grid computing implementation in Pakistan. Using PERN would enable us to reach our target audience (universities and research institutions) in a cost effective and quick manner. On the other hand, PERN currently support lower bandwidth for most scientific and related grid applications. In future though, we expect the university to backbone connectivity to increase to at least 10 Mbps to make this network more useful.

This would be possible if Optical Fiber is used in place of the current E1 lines and the recurring costs would be low since the international bandwidth should not necessary equal the link speed.

The proposed architecture making use of PERN network is displayed in Figure 13. The architecture can scale with the locations and users. The clients can access the grid resources by various methods, namely web browser, Globus or service client tools. The main access is via Portal Server which makes use of the Open Grid Computing Environment (OGCE). This enables users to access most resources without setting up the client tools. Additionally, the resources can be accessed from anywhere around the world since web browsers are universally available.

As discussed earlier, Grid Certification Authorities (Grid CA) is one of the main components of a grid. It is a common practice to have the grid certification authority accessible only to the grid participants and cater only to the need of the grid and that is why these are kept separate from the PKI Certification and Registration Authorities. According to this, we ought to have a "Pakistan Grid Certification Authority" and a Registration Authority. These will be responsible for giving out user and host certificates after verifying the credentials of the users.

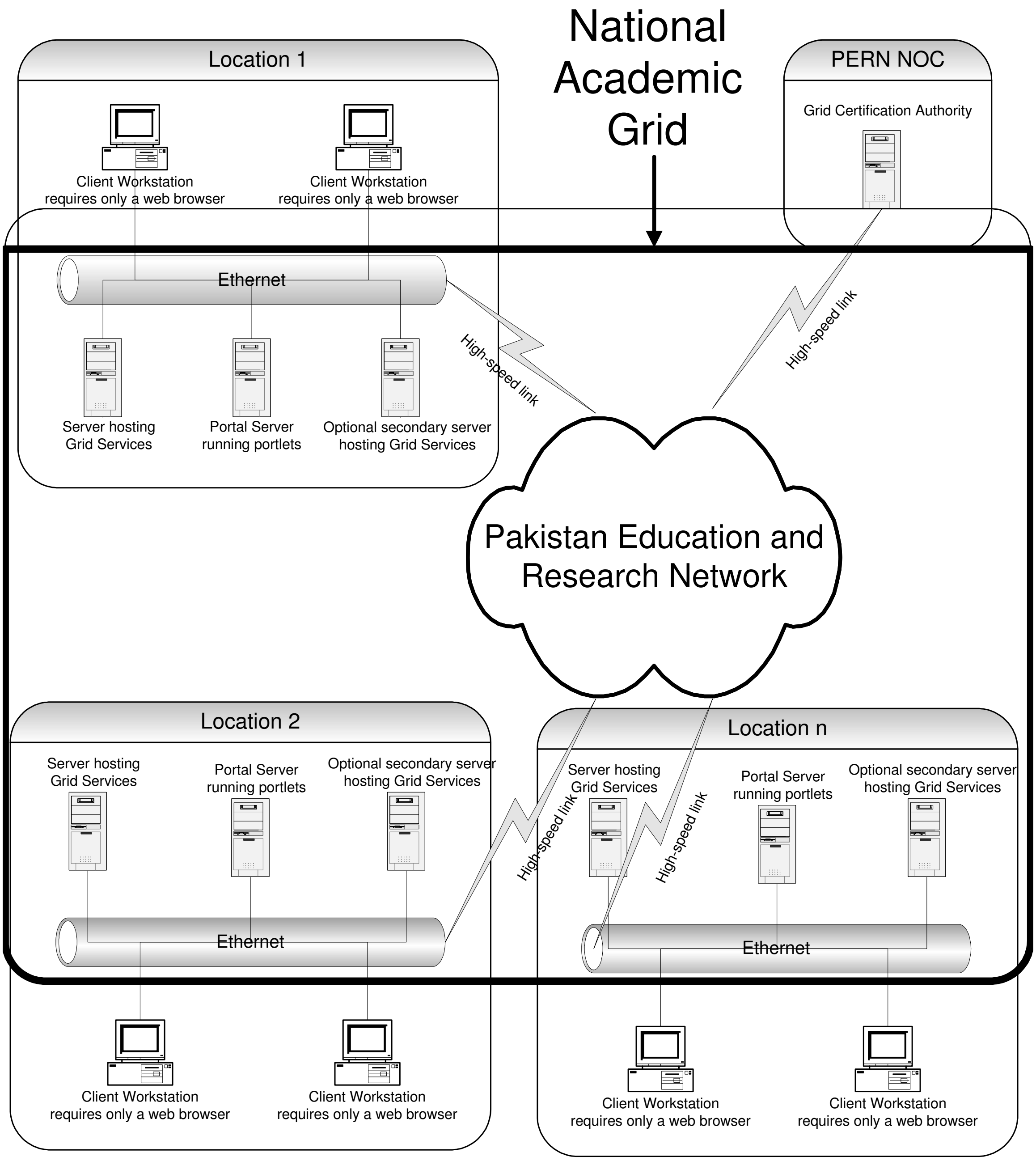

Figure 13: National Academic Grid Architecture

3.4 Proposed Services

In order to properly utilize such a large distributed structure, we need to have some useful services, without which a grid would not be complete. The following are some of proposed services, each of which includes many sub-services. The discussion will include the service introduction, interface, brief implementation comments, and some user interface details. A typical issue, faced by various institutions, of locating some grid resource or performing particular action on a grid, thus becomes as simple as finding the right grid service or writing one.

All communication between the client and the server would be in XML, while the portal server would transform the XML using XSL and transmit the result as a HTML page. The following is a list of some common services and user interface features.

Common Services:

- **Security services:** Services and operations to handle authentication, authorization and other common security issues.
- **Metadata and replica management service:** Maintains metadata about services and support in replication of data amongst various locations.
- **QoS management and network load balancing service:** The service helps other services in maintaining quality of service and balancing load when multiple distributed services are available.

Common User Interfaces:

- **Catalog pages:** Search pages for normal and advanced search options. The results are transformed from xml and displayed on the screen for the user along with the new search, reserve and download options.
- **Document management:** The upload and download of documents will make use of the Grid FTP service.

3.4.1 Learning Services

Many universities have been running distance learning and teaching and research collaboration programs for a long time. Grid computing has the ability to enhance this collaboration to a great extent. It can be used as a Learning Management System (LMS) with much more capabilities than a traditional LMS. Uses include curricula development, virtual lectures, quizzes, experiments and collaboration. Meetings can be scheduled as well between teachers and students and guest lecturers can participate as well. The following services are derived from [30].

Services / Operations:

- **Program / course catalogs:** Search the programs and courses offered for a particular course or program. Search criterion and results will be transmitted in XML format.
- **Course materials catalogs:** Search for course materials (lectures, handouts, etc) and presents the list of matching documents.
- **Document management service:** View a given course material document or upload a new document (lecture, handout, assignment) in the course material
- **Registration service:** Manage registration and examination details
- **Student performance / assessment service:** Grading and evaluation of assignment, quizzes, etc as well as the assessment of the teachers.
- **Quiz / homework services:** Submission of quizzes and homework assignments as well as results. Manages deadlines.
- **Meeting / class scheduling service:** Schedule any meeting with the teacher, or teacher's assistant. Also receive and update the latest class schedule.

3.4.2 Digital Library for Documents

A digital library service should include a wide range of sub-services ranging from authoring to editing and searching documents. In most institutions, there are a lot of documents like lectures, research papers, reports and others. Easy access and sharing of these resources has always been problematic. With such a grid service in place, we can easily author the documents as well as supporting fast searching, replication and location transparency.

Services / Operations:

- **Catalog / search service:** Search and catalog service to search for documents.
- **Document management service:** Service supporting creating, editing, changing properties and archiving documents.
- **Document rights management:** Manages and implements rights on documents and categories.
- **View document:** View or download the document in convenient formats from the closest available location.

3.4.3 Library Catalog Service

Most of the universities in Pakistan have a limited number of books in Libraries and thus the students face a hard time if the book is not available in the local library. A shared library catalog will help the student to read some very important text by borrowing it from the other library.

Services / Operations:

- **Catalog / Search Service**: Given a book title query for the book return result in xml. Contains simple, detailed and peer query and catalog details.
- **Manage book subscriptions:** Reissue, reserve or calculate fine on books.
- **Request scanned pages:** Service to request electronic copies of some of the pages of the books.

User Interface:

- **Status pages:** Two status pages to look at the books currently issued on user's card and then, either re-issue it or calculate the fine on it.

3.4.4 Multimedia Services

The trend during the last decade has been towards the increasing use of videos and multimedia aids for business, learning and research. Grid can also be used for this purpose. In an educational setting, such a service is used in conjunction with the learning and digital library services as described above.

Services / Operations:

- **Catalogs / search service:** Service to manage catalog and search the contents.
- **Stream / download contents:** Stream / download contents from the closest available location
- **Content management service:** Service supporting Add, Update, archive media content.
- **Content rights management service:** Manage the rights of the media contents.
- **Manage categories:** Manage the categories of the media contents.

3.4.5 Scientific Data Services

Most scientific experiments are producing a large amount of data. Grid services are required to effectively organize such a vast amount of information distributed across the globe. Also, searching and processing such information requires large amount of processing power. Grid is one of the most promising technologies which can be used for this purpose.

Services / Operations:

- **Catalogs / search service:** Locate the content for the users
- **Data management:** Allows users to create, manage the way data is stored. Provides single view of multiple and distributed specialized types of storage like files, databases, etc.
- **Data query and analysis:** Service to query and analyze large volume of distributed data. This should also allow users to submit data to specialized scientific equipments available on the grid for analysis.

3.4.6 Software Development Services

A lot of software development work is being outsourced these days. Similarly, a lot of projects, notably the open source ones have large distributed teams working from around the world. Internet has made such collaboration pretty effective, but a grid would make it much more powerful and useful. Examples include grid-enabled design tools, source code repository, and document management.

Services / Operations:

- **Version control service:** A version control service to control the software development process and consistency and availability of source code.
- **Distributed build:** Enable distributed builds of large applications in order to save time.
- **Software design service:** Service to enable collaborative design of software. The service could be made advanced enough to help two different teams present at different locations to work on two aspects of the design.

Chapter 4. Prototype Grid: Implementation & Experiences

In order to demonstrate the feasibility of such an environment, a prototype grid was setup within the campuses of our university in Karachi. Some sample grid services were also implemented and deployed on top of that grid.

4.1 The Architecture

Although the PAREGIN should use PERN, the initial prototype uses LAN. The architecture of the prototype grid is similar to the proposed one except the use of Local Area Network to simulate the PERN-based implementation. Replacing LAN with PERN will not require any major change in the architecture; rather the changes only need to be made at the network layer. This was tested using public IP addresses of different ISPs and all traffic passed through routers. Figure 14 shows the architecture of the prototype implemented.

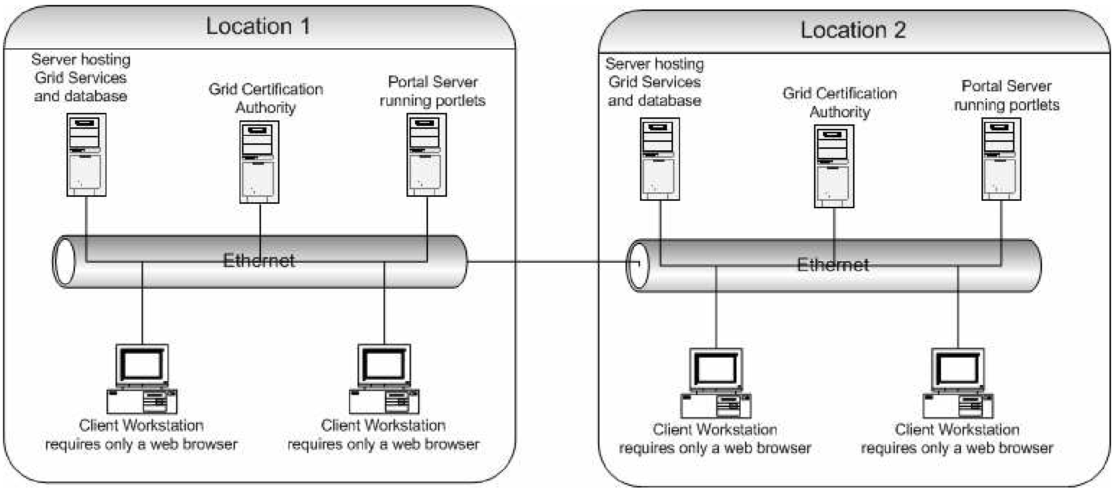

Figure 14: Prototype Architecture

4.2 Services for prototype grid

The following three main services were identified to demonstrate the usefulness of grid Infrastructure.

4.2.1 Shared Library Catalog

Most of the universities in Pakistan have a limited number of books in Libraries and thus the students face a hard time if the book is not available in the local library. A shared library catalog will help the student to read some very important text by borrowing it from the other library. The service need to be able to help the users who need to query the database and also to replicate the data to remote locations. All operations need to be transparent.

4.2.2 Document Access Service

In any given organization, there are a lot of documents, memos, notices, going on. Also, teams working on collaborative projects and teachers teaching the same courses often want to share documents and large experimental or analytical data. All of these scenarios demand sharing of documents in a secure manner from various locations. The current plan is to make a document access service to search and access the documents available on the grid. Additionally, service could also handle replication of data from one location to another. The future versions would handle uploading and managing per-location rights management.

4.2.3 Shared electronics books

There are a large number of electronic books coming with the books bought by us. All these electronic books plus those freely available on the internet can be shared in ways similar to the documents.

4.3 The Application Service Implementation

The middleware used in this prototype grid implementation is based on Globus Toolkit 3 API, libraries and services. The standard operating system of all the nodes is Linux with JRE 1.4.2.

4.3.1 The Application Service Architecture

A grid infrastructure has been devised for the development of various grid applications on this grid. Since the major focus of such architecture will be on data-related applications of the grid, we emphasize on data replication, consistency and transparency. Such replication and indexing is done via the Globus toolkit. The toolkit also provides mechanism to find the most useful copy of the data. The first phase of development of such an application would be to search the data in the local pool of data including the replicated data. The second phase searches the relevant data at other locations. This is shown in Figure 15.

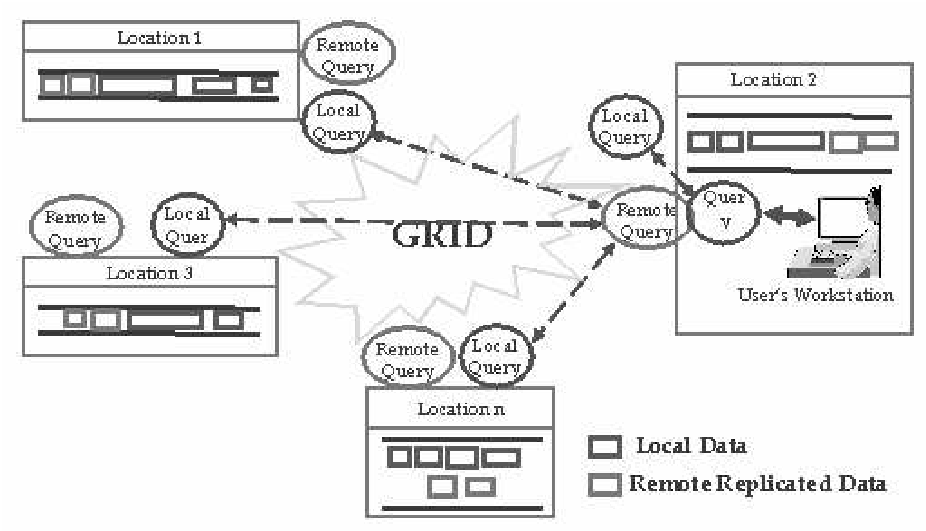

Figure 15: Grid Infrastructure for Data-Intensive Applications

The programming model of a grid service is similar to that of web service [31]. We first have to make a schema describing the operations to be supported by the service. It is then compiled to form a stub class which is used to implement clients and the actual service conforming to the schema. The grid service schema is a XML document and describes the operations, there parameters and the data types used by them.

The second step is to actually code the service conforming to the schema. The service queries the local database for locally available and replicated information. It then calls on the remote services to get results from them, finally combining the results and sending them to client. In order to make the design scalable, any number of services can be mentioned in the configuration file which resides on the local machine. The service, each time it is invoked, reads the list of available remote services and call hem automatically. The database machine configuration can also be modified using the configuration file.

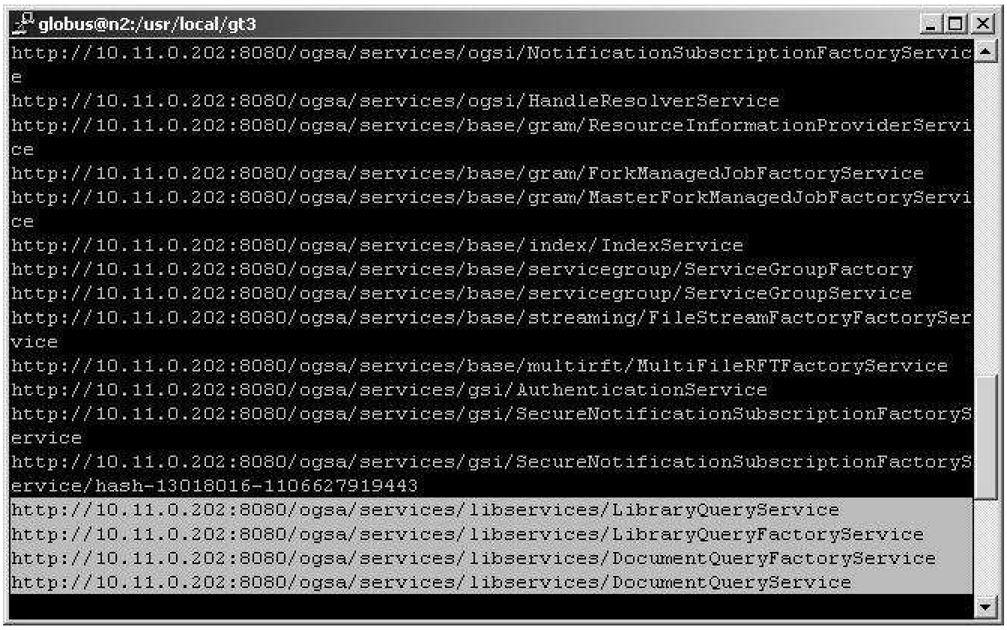

Figure 16: A snapshot showing the running Globus implementation along with two application services on a Linux machine

4.3.2 Application Service Implementation

The Query Services needs at least the following operations:

- Query – query the local, replicated and remote data for some book
- peerQuery – handles query from a peer service
- getCatalogDetails – provides details of catalog

All communication to and from these operations would be in the form of XML strings, while all file transfer operations are conducted using the Grid FTP.

Abstract from the service descriptor (gwsdl file) is shown in Table 3**Error! Reference source not found.**. Some portions have been striped off to shorten the length of the file. While a sample XML result is shown in 4.

```
<?xml version="1.0" encoding="UTF-8"?>
<definitions name="LibraryQueryService"
      targetNamespace="http://khi.nu.edu.pk/services/namespaces/2004
/01/libservices/LibraryQueryService"
      xmlns:tns="http://khi.nu.edu.pk/services/namespaces/2004/01/li
bservices/LibraryQueryService"
      xmlns:ogsi="http://www.gridforum.org/namespaces/2003/03/OGSI"
      xmlns:gwsdl="http://www.gridforum.org/namespaces/2003/03/gridW
SDLExtensions"
      xmlns:xsd="http://www.w3.org/2001/XMLSchema"
      xmlns="http://schemas.xmlsoap.org/wsdl/">

<import location="../../ogsi/ogsi.gwsdl"
      namespace="http://www.gridforum.org/namespaces/2003/03/OGSI"/>

<types>
<xsd:schema
targetNamespace="http://khi.nu.edu.pk/services/namespaces/2004/01/li
bservices/LibraryQueryService"
      attributeFormDefault="qualified"
      elementFormDefault="qualified"
      xmlns="http://www.w3.org/2001/XMLSchema">

      <xsd:element name="query">
            <xsd:complexType>
                  <xsd:sequence>
                        <xsd:element name="value"
type="xsd:string"/>
                  </xsd:sequence>
            </xsd:complexType>
      </xsd:element>
      <xsd:element name="queryResponse">
            <xsd:complexType>
                  <xsd:sequence>
                        <xsd:element name="value"
type="xsd:string"/>
                  </xsd:sequence>
            </xsd:complexType>
      </xsd:element>
…
</xsd:schema>
</types>
<message name="QueryInputMessage">
      <part name="parameters" element="tns:query"/>
</message>
<message name="QueryOutputMessage">
      <part name="parameters" element="tns:queryResponse"/>
```

```
</message>
…
<gwsdl:portType name="LibraryQueryPortType"
extends="ogsi:GridService">
      <operation name="query">
            <input message="tns:QueryInputMessage"/>
            <output message="tns:QueryOutputMessage"/>
            <fault name="Fault" message="ogsi:FaultMessage"/>
      </operation>
      <operation name="peerQuery">
            <input message="tns:PeerQueryInputMessage"/>
            <output message="tns:PeerQueryOutputMessage"/>
            <fault name="Fault" message="ogsi:FaultMessage"/>
      </operation>
      <operation name="detailedQuery">
            <input message="tns:DQInputMessage"/>
            <output message="tns:DQOutputMessage"/>
            <fault name="Fault" message="ogsi:FaultMessage"/>
      </operation>
      <operation name="getCatalogDetails">
            <input message="tns:GetCatDetailsInputMessage"/>
            <output message="tns:GetCatDetailsOutputMessage"/>
            <fault name="Fault" message="ogsi:FaultMessage"/>
      </operation>
</gwsdl:portType>

</definitions>
```

Table 3: Library query service descriptor file


```
<?xml version="1.0"?>
<queryresult query="toefl">
<resultset server="local">
<result xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"><row><Title>TOEFL</Title><SubTitle>Test of English
as a Foregh Languages</SubTitle><ISBN
xsi:nil="true"/><Publisher>CBS
Publishers</Publisher><Author1>Rudman,Jack.</Author1><Year_Of_
Pub>1994</Year_Of_Pub><Edition xsi:nil="true"/></row></result>
</resultset>

<resultset
server="http://n2:8080/ogsa/services/libservices/LibraryQueryS
ervice">

</resultset>
</queryresult>
```


Table 4: Sample Result of LibQueryService

4.3.3 Grid Portal Implementation

Setting up the grid client tools is a hectic and lengthy process. So, we need to provide a user interface to grid users that is easier to setup and accessible from many different places including researcher's home and may be from even Internet. Thus, we need to implement a grid portal. After evaluating some portal toolkits and APIs, the Open Grid Computing Environments (OGCE) was selected. OGCE combines the Apache Tomcat Servlet Container and Apache Jetspeed Portlet Engine as well as portlets from various sources. Thus, a user need only a web browser to access the grid resources.

4.3.4 Using the services via the Portal

Two JSP Portlets, a type of portlet, have been developed (as shown in the figures below) to provide users ability to search and download resources available on the grid using their web browser. The portlets get the query from the user and invokes the service with the query as parameter. The results, in XML are transformed to HTML using XSL style sheet and sent to the user's browser for display. The portlet has been coded so that even if the main grid service machine is down; it will get the query results from secondary or any remote machines whichever is available.

The following figures show the portal being accessed by a web browser running on a Windows machine. The browser is accessing the OGCE portal running on a dedicated portal server, while the portal server is accessing the grid services on behalf of the user logged in.

The setup also includes myProxy server, which provides the service of keeping track of the users' credentials in the form of an encrypted certificate. The OGCE portlet setup is integrated with myProxy, so the user just needs to log on to the server once in a week to refresh his/her certificate and can then access the grid services from anywhere in the world.

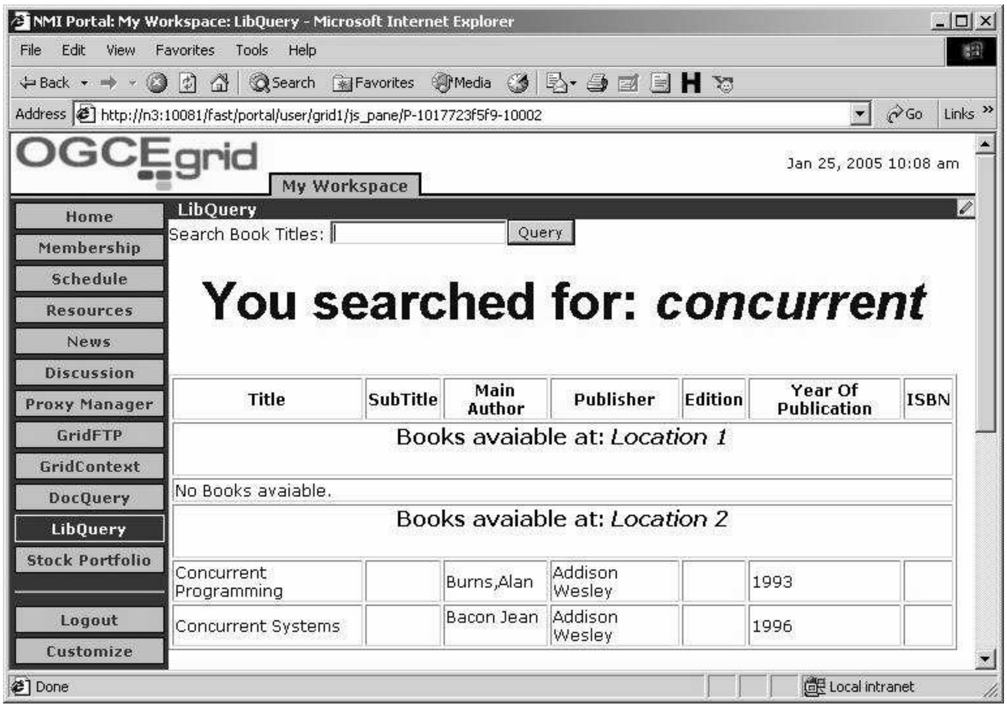

Figure 17: The Library Query Service portlet

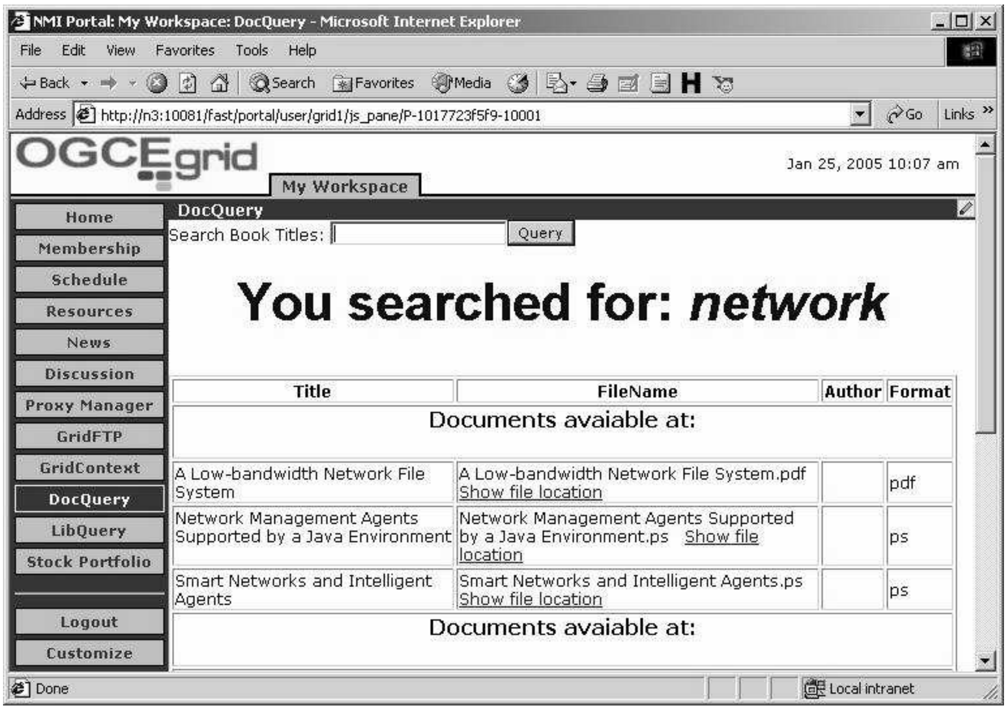

Figure 18: The Document Query Service Portlet

The grid portal come with a number of grid and web portlets preloaded. One of the portlets is the GridFTP Portlet using which one user can view and download the files to which he has been granted access. Grid FTP Portlet lets you use the Grid FTP service on the remote machine using your web browser. The only requirement is the presence of a valid proxy certificate.

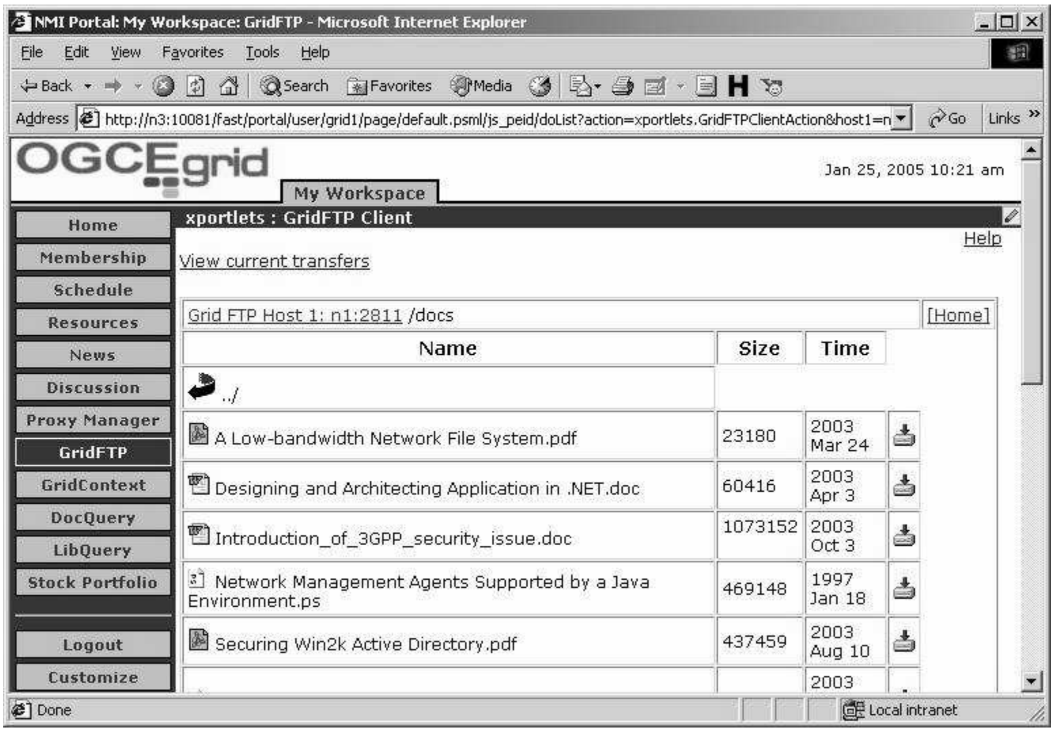

Figure 19: GridFTP client portlet

Chapter 5. Evaluation of the Grid Implementation

We evaluated the prototype grid and the proposed architecture in order to understand and identify its advantages and limitations.

5.1 Reuse of existing resources

The solution focuses on making heavy use of the existing infrastructure. Instead of proposing a new network backbone connecting all the institutions, PAREGIN will make use of the PERN in order to lower the cost and the effort involved. Similarly, the universities do not need to have dedicated new hardware for PAREGIN, instead the service providers simply need to deploy the grid services on any of their current systems, while the service users can use a simple web browser, available on all computers.

5.2 Ease of deployment

The deployment of grid clients and services is a difficult process especially if you have a large number of institutions involved. In order to lessen the work involved, various scripts were made so as to make the deployment of grid middleware and various services easy and fast. The client side mainly consists of browser based users, which removes any work to set up a client. The user is only required to logon once a week to refresh his / her certificate and then can use a web browser from anywhere in the network to access the grid resources.

5.3 Cost Feasibility

The basic building block of a grid is the Network Infrastructure. Since we have proposed to make use of the Pakistan Education and Research Network (PERN), the cost of setting up such a grid would be minimized. Definitely this would affect the performance of the Applications that we would be running on top of the grid. Another benefit of using PERN would be the reach of PAREGIN to far flung Universities and Research Institutions.

5.4 Usability Analysis

All the proposed services would now be analyzed for the benefits and any probable issues.

5.4.1 Learning Services

Just like the Learning Management Systems (LMS) that we currently use, a grid based learning service would be helpful in inter-institution and intra-institution communication and collaboration. This would be extremely helpful in devising course outlines for newer courses, collaborative projects and research.

5.4.2 Digital Library for Documents

Such a service can be used for universal access to authoring, searching, and reading of comparing various types of documents.

5.4.3 Library Catalog Service

The limitation of availability of latest books in our library can be overcome if we have a combined catalogs and inter-library loan facility in place. The first step is to have an easy to use nation-wide book catalogue which can be very easily built on top of a grid as a Library Query Service.

5.4.4 Multimedia Services

Multimedia resources are now being used in areas ranging from formal education, trainings, and tutorials to real-life examples. Internet has given a huge boast to the availability of such resources. Grid can take these kinds of services a step ahead because of the ease of integration of these.

5.4.5 Scientific Data Services

The huge amount of data being produced by the various scientific experiments need to be optimally stored and processed. Grid is one of the most ideal solutions for this purpose and many teams all over the world are working on exploiting grid for this purpose. Such a service will consist of a lot of services, including, data management services, and processor sharing services.

5.5 Scalable Design

The design has the ability to change the resources and infrastructure capability at any moment. The resource management mechanism supported by Globus has been used and it takes care of resource registration, discovery and usage in a transparent manner. Similarly the network infrastructure could easily be upgraded to a better one without any affect on the existing services.

The scalability of individual services is the responsibility of the person writing the service while the support for this is present in the architecture.

5.6 Performance Analysis

Since only some of the services were implemented in the prototype grid, we can examine the performance aspect of this service briefly.

All of these services are built on top of a common architecture which has been described in the previous chapter. The client would request these services using the grid portal front-end. The portal services would contact the grid services using the user proxy credentials which are retrieved automatically. The user will then use these services using the easy to use browser interface. In order to search, the user can enter a word in the search box and search for it. If this is a document, user will be presented with a link for downloading it along with the search results. All file downloads are performed using the easy to use interface for Grid-FTP.

5.7 Technologies used

The solution makes use of emerging technologies like XML, XSL, WSDL, Grid Portals, and the Grid Services. These and many other technologies make the whole solution highly configurable and easy to program the proposed and any new set of services.

5.8 Institution's Benefits

The institutions involved will benefit from PAREGIN in a number of ways as discussed in Chapter 3. The following lists some of the important benefits an institution will gain:

- Access to expensive resources of other institutions
- Collaborative and distant learning
- Increased collaboration amongst teams working on similar problem at various institutions

5.9 Technical Issues

Some of the technical problems might arise when implementing grid computing in the next few years. These are expected to be solved as more and more research is conducted.

- **Programming Tools:** We require grid specific programming/problem solving tools for developers to write for grid platforms easily and efficiently.
- **Service Quality:** Ensuring the Quality of Service is still an unresolved issue. If the resource needed are not available or do not function as desired than the business case for grid computing is lost. Defining a Service Level Agreement would also be a problem as there is currently little understanding of such issues.

- **Resource Management:** According to [33], grid must be able to quickly ascertain what resources are available on any computer that joins it. The grid shouldn't be bogged down by a slow or outdated system. Avoiding the "least common denominator" problem is somewhat high on the technical "to-do" list.
- **Security:** Under no circumstances would anyone want free and open access to the computing resources. Proper authentication and authorization methods are an essential need.
- **Policy Enforcement and Representation:** A proper framework for grid policy is required using which an organization can make and enforce the necessary policies.
- **Failure detection and recovery:** Failure detection and recovery is of even more importance in grid as multiple stakeholders are involved.

5.10 Non-technical issues

The main problem in implementing such a grid would be the following:

- The network infrastructure required for any useful grid application is a lot. Absence of such a high-speed network might affect the overall usefulness of the grid. This has been discussed in detail in earlier sections.
- The administration of most universities would not readily allow the sharing of resources.
- No major research activity is going on in the universities, let alone some collaborative activity.

Chapter 6. Conclusion and Future Work

The WWW and the Internet allowed limited collaboration in the form of groupware and collaboratories. With the advent of grid, academicians have found new and improved ways to collaborate with each other. A grid enables participants to share resources with each other at the same time controlling the access to these resources.

The thesis presented design and analysis of a collaborative grid framework within a developing country like Pakistan along with a prototype implementation. Many teams have been working on implementing grid computing in developed countries, but Pakistan being a developing country has completely different scenario. Lack of any high-speed nation wide network is one issue and the absence of collaboration amongst local institutions is another.

The recent PERN project is expected to connect all the educational and research institutions of the country together is a good start. We need to have some spare bandwidth (5Mbps would be a good starting point) so that it can be utilized for collaborative applications. Once such a setup is in place we can implement a proper academic collaborative grid and collaboration amongst the institution would automatically start. The PERN network could further be developed with the help of industry partners and at least be brought as close to the current national, organizational and desktop link ratios of 10:1:0.1 Gbps.

The use of open source tools and APIs at he middleware, and backend with web browsers at the front end would mean that the software needed for such a large scale deployment would be community driven and patches and updates would be released quite frequently.

The experience from the implementation of the services and portlets, led us to belief that many useful services could be made and deployed on top of such a grid by the students and researchers of the participating institutions. Some starting applications could be a complete document management service, scientific data sharing and analysis services and many others.

6.1 Future Work and issues in Grid Implementation

6.1.1 Issues in Grid Implementation

A grid of such a magnitude would be having a lot of problems including:

- Lack of reliable and fault-tolerant network infrastructure.
- Not many resources would be shared by Universities.
- No current collaboration underway
- Lack of Programming Tools
- Ensuring the Quality of Service is still an unresolved issue
- Lack of proper Resource Management Tools
- Security Considerations would also hinder full blown resource sharing
- Lack of a proper framework for grid policy.

These and many other issues had been declared in depth in chapter 5.

6.1.2 Wireless interfaces to the grid

Just like a user can use a web browser, she can also use her WAP, GPRS or Wi-Fi enabled wireless device to search the grid resources. Portlets can be used here as well to provide easy to use user interface.

6.1.3 Technical Challenges for Wireless Grids

Definitely, wireless grids have many additional technical challenges and many of the solutions we use in wired grids are not useful for the wireless ones. A number of technical challenges are listed in [32] which are summarized below:

- **Dynamic Configurability:** Wireless grids must offer high degree of temporal flexibility. This is because the mobile nature of the grid components results in topologies that change frequently over time.
- **Routing Plasticity:** Efficient routing protocols are required to address the power limitation of the end devices along with the consideration for stable wireless connectivity, route optimization and efficient use of the limited bandwidth.
- **Discovery Semantics and Protocols:** Service description protocols are needed to describe the services provided by various components of the wireless grid. Once these services are published, a discovery protocol is needed to map the mobile resources to the services.
- **Security:** Wireless grids must incorporate effective mechanisms to provide adequate security and privacy to end-users.
- **Policy Management:** Wireless grids must incorporate appropriate policies to support both pre-conceived and agile interactions among the relevant sets of users.

References

[1] Foster, I., Kesselman, C. (eds.), *"The Grid: Blueprint for a Future Computing Infrastructure,"* Morgan Kaufmann, 1998.

[2] Tugend, T., *"UCLA to be the first station in Nationwide Computer Network,"* UCLA Press Release, http://www.lk.cs.ucla.edu/LK/Bib/REPORT/press.html, July 1969.

[3] Smarr, L., Catlett, C., *"Metacomputing*," Communications of the ACM, 35 (6) 44-52, 1992.

[4] Foster, I., Kesselman, C., *"Globus: a Metacomputing Infrastructure Toolkit,*" International Journal of Supercomputer Applications. 11 (2), pp. 115-128, 1997.

[5] Foster, I., *"Computational Grids,*" In: Foster, I., Kesselman, C. (eds.), *"The Grid: Blueprint for a Future Computing Infrastructure,"* Morgan Kaufmann, pp. 15-52, 1998.

[6] Foster, I., Kesselman, C., Tuecke, S., *"The Anatomy of the Grid: Enabling Scalable Virtual Organizations,"* International J. Supercomputer Applications, 15(3), 2001.

[7] Ong, S. H., *"Grid Computing: Business Policy and Implications,"* Master's Thesis, MIT, Cambridge, MA, 2003.

[8] Foster, I., *"What Is the Grid? A Three Point Checklist"*, Grid Today, 1 (6), 2002.

[9] Roure, D.D., et al., *"The Evolution of the Grid,"* In: Berman, F., Fox, G., Hey, T. (eds.), *"Grid Computing: Making The Global Infrastructure a Reality,"* John Wiley & Sons, pp. 65-100, 2003.

[10] Berman, F., Fox, G., Hey, T. *"The Grid: past, present and future,"* In: Berman, F., Fox, G., Hey, T. (eds.), *"Grid Computing: Making The Global Infrastructure a Reality,"* John Wiley & Sons, pp. 9-50, 2003.

[11] Smarr, L., *"Grid in Context"*. In Foster, I., Kesselman, C. (eds.) *"The Grid 2: Blueprint for a New Computing Infrastructure,"* Morgan Kaufmann, pp. 8-12, 2003.

[12] Object Management Group, *"Common Object Request Broker: Architecture and Specification,"* Revision 2.2, Document 96.03.04, 1998.

[13] Arnold, K., et al., *"The Jini Specification,"* Addison-Wesley, 1999. See also Jini website: http://www.sun.com/jini.

[14] Clark, D., *"Face-to-Face with Peer-to-Peer Networking,"* Computer, Vol. 34, No. 1, pp. 18-21, January 2001.

[15] Fox, G., et al., *"Peer-to-peer Grids,"* In: Berman F., Fox G., Hey T. (eds.), *"Grid Computing: Making The Global Infrastructure a Reality,"* John Wiley & Sons, pp. 471-489, 2003.

[16] TeraGrid Project, http://www.teragrid.org/.

[17] Pakistan Education and Research Network (PERN) Website, http://pern.edu.pk

[18] PERN Architecture, http://www.pern.edu.pk/home/network_archeticture.htm

[19] Papazoglou, M., Georgakapoulos, D., *"Service oriented computing,"* Communications of the ACM, 46(10):24–28, October 2003.

[20] Foster, I., Kesselman, C., Nick, J.M., Tuecke, S., *"The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration, "*In: Berman F., Fox G., Hey T. (eds.), *"Grid Computing: Making The Global Infrastructure a Reality,"* John Wiley & Sons, pp. 217-250, 2003; and http://www.globus.org/research/papers/ogsa.pdf..

[21] Unger, J., Haynos, M., *"A visual tour of Open Grid Service Architecture,"* IBM Developerworks, http://www-106.ibm.com/developerworks/grid/library/gr-visual/, August 2003.

[22] Tuecke, S., et al., *"Open Grid Services Infrastructure (OGSI) version 1.0,"* Global Grid Forum, http://www.ggf.org/ogsi-wg, June 2003.

[23] Gentzsch, W., *"Grid Computing: A New Technology for the Advanced Web,"* White Paper, Sun Microsystems, Inc., Palo Alto, CA, 2001.

[24] Agarwal, A., Norman, D. O., Gupta, A., *"Wireless Grids: Approaches, Architectures and Technical Challenges,"* Working Paper 4459-04, MIT Sloan School of Management, January 2004.

[25] Platform Computing, *"Understanding Grid Computing,"* Whitepaper, April 2002.

[26] Platform Computing, *"The Evolution of Grid: The Three Stages of Grid Computing,"* http://www.platform.com/grid/evolution.asp.

[27] McRobbie, M.A., et al., *"A Global Terabit Research Network,"* http://www.gtrn.net/global.pdf.

[28] Tiang, H., *"Grid Computing as an Integrating Force in Virtual Enterprises,"* Master's Thesis, MIT, Cambridge, MA, 2003.

[29] Maldonado, M.F., *"Grid computing in higher education: Trends, values and offerings,"* IBM Corporation, December 2004.

[30] Fox, G., et al., *"Education and the enterprise with the Grid"* In: Berman F., Fox G., Hey T. (eds.), *"Grid Computing: Making The Global Infrastructure a Reality,"* John Wiley & Sons, pp. 471-489, 2003.

[31] Sotomayor, B., *"The Globus Toolkit 3 Programmer's Tutorial"*, Avaiable from http://www.casa-sotomayor.net/gt3-tutorial, 2004

[32] Agarwal A., Gupta A., Norman D. O., *"Self-Configuration and Self-Administration of wireless grids,"* Working Paper 4460-04, MIT Sloan School of Management, January 2004.

[33] IBM Developerworks, *"Grid computing: Conceptual flyover for developers,"* http://www-106.ibm.com/developerworks/library/gr-fly.html, May 2003

Appendix I: Some Selected Grid projects

Name	Major Focus & Website
Access Grid	Create & deploy group collaboration systems using commodity technologies Grid test bed linking IBM laboratories http://www.accessgrid.org
DOE Science Grid	Create operational Grid providing access to resources & applications at U.S. DOE science laboratories & partner universities http://sciencegrid.org
Earth System Grid	Delivery and analysis of large climate model datasets for the climate research community http://earthsystemgrid.org
EU DataGrid	Create & apply an operational grid for applications in high energy physics, environmental science, bioinformatics http://eu-datagrid.org
Globus Project	Research on grid technologies; development and support of Globus Toolkit ™; application and deployment http://globus.org
Gridbus	A toolkit for service-oriented computing. http://www.gridbus.org
Grid Physics Network	Technology R&D for data analysis in physics experiments: ATLAS, CMS, LIGO, SDSS http://griphyn.org
International Virtual Data Grid Laboratory	Create international Data Grid to enable large-scale experimentation on Grid technologies & applications http://ivdgl.org
NASA Information Power Grid	Create and apply a production grid for aero sciences and other NASA missions http://ipg.nasa.gov
Particle Physics Data Grid	Create and apply production grids for data analysis in high energy and nuclear physics experiments http://ppdg.net
TeraGrid	US science infrastructure linking four major resource sites at 40 Gbps http://teragrid.org
UK eScience Grid	The use of grid for constructing Science applications. http://www.nesc.ac.uk

Appendix II: Status of PERN Implementation

S#	Name of the University	Last Mile	Planned B.W	Provisioned B.W
Islamabad Region				
1	Air University, Islamabad	DXX	2Mbps	128Kbps
2	Allama Iqbal Open University, Islamabad	SDH	3Mbps	4Mbps
3	Bahria University, Islamabad	DXX	2Mbps	1Mbps
4	COMSATS Institute of Information Technology, Islamabad	DDX+Radio	4Mbps	4Mbps
5	Fatima Jinnah Women University, Rawalpindi	DXX	2Mbps	1Mbps
6	Ghulam Ishaq Khan Institute of ES & Tech, Swabi	DDX+Radio	4Mbps	1Mbps
7	Gomal University, D. I. khan	Radio	2Mbps	1Mbps
8	Higher Education Commission, Islamabad	DRS	4Mbps	4Mbps
9	International Institute of Space Technology, Islamabad	DRS	2Mbps	2Mbps
10	International Islamic University, Islamabad	DXX	2Mbps	1Mbps
11	National Defense College, Islamabad	DRS	2Mbps	2Mbps
12	National University of Computer & Emerging Sciences (FAST), Islamabad	DXX+Radio	4Mbps	1Mbps
13	National University of Modren Languages, Islamabad	DXX	2Mbps	1Mbps
14	National University of Science & Technology, Rawalpindi	DDX+Radio	4Mbps	2Mbps
15	NWFP University of Agriculture, Peshawar	DXX	2Mbps	512Kbps
16	Pakistan Institute of Engineering & Applied Sciences, Nilore	DRS	3Mbps	4Mbps
17	Pakistan Military Academy, Kakool	DRS	2Mbps	2Mbps
18	Peshawar University, Peshawar	SDH	3Mbps	4Mbps
19	Quaid-i-Azam University, Islamabad	SDH	2Mbps	4Mbps
20	University of Arid Agriculture, Rawalpindi	Radio	2Mbps	512Kbps
21	University of Azad Jammu & Kashmir, Muzafar Abad	DXX	2Mbps	2Mbps
22	University of Engineering & Technology, Peshawar	SDH	4Mbps	4Mbps
23	University of Engineering & Technology, Khuzdar	VSAT	2Mbps	2Mbps
24	University of Engineering & Technology, Taxila	DRS	4Mbps	4Mbps
Lahore Region				
25	Bahauddin Zakariya University, Multan	SDH	2Mbps	2Mbps
26	Government College University, Lahore	SDH	2Mbps	2Mbps
27	Islamia University, Bahawalpur	DRS	2Mbps	2Mbps
28	Lahore College for Women University, Lahore	DXX	2Mbps	2Mbps
29	Lahore School of Economics, Lahore	DRS	4Mbps	4Mbps
30	Lahore University of Management Sciences (LUMS), Lahore	SDH	4Mbps	256Kbps
31	National College of Arts, Lahore	DXX	2Mbps	1Mbps
32	National University of Computer & Emerging Sciences (FAST), Lahore	DXX+Radio	4Mbps	2Mbps
33	Pakistan Administrative staff College, Lahore	DXX	2Mbps	2Mbps
34	Punjab University, Lahore	SDH	4Mbps	4Mbps

35	University of Agriculture, Faisalabad	DXX+Radio	3Mbps	3Mbps
36	University of Engineering & Technology, Lahore	SDH	4Mbps	4Mbps
Karachi Region				
37	Agha Khan University, Karachi	SDH	3Mbps	3Mbps
38	Balochistan University of Information Technology & Management Sciences, Quetta	DRS	2Mbps	2Mbps
39	College of Physicians & Surgeons Pakistan, Karachi	DXX+Radio	3Mbps	1Mbps
40	Command & Staff College, Quetta	DRS	2Mbps	2Mbps
41	Hamdard University, Karachi	DRS	3Mbps	2Mbps
42	HEJ Institute of Chemistry, International Center for Chemical Sciences, Karachi	DRS	4Mbps	4Mbps
43	Institute of Business Administration, Karachi	DRS	2Mbps	2Mbps
44	Karachi Institute of Information Technology, Karachi	DXX	2Mbps	2Mbps
45	Mehran University of Engineering & Technology, Jamshoro	SDH	4Mbps	4Mbps
46	National University of Computer & Emerging Sciences (FAST), Karachi	DRS	4Mbps	4Mbps
47	NED University of Engineering & Technology, Karachi	SDH	4Mbps	2Mbps
48	Pak Naval Academy, Karachi	DRS	2Mbps	2Mbps
49	Quaid-e-Awan University of Engineering, Science & Technology, Nawabshah	DXX+Radio	4Mbps	128Kbps
50	Shah Abdul Latif University, Khairpur	DRS	2Mbps	2Mbps
51	Shaheed Zulfikar Ali Bhutto Institute of Science & Technology (SZABIST), Karachi	DXX	2Mbps	2Mbps
52	Sindh Agriculture University, Tandojam	SDH	2Mbps	2Mbps
53	Sir Syed University of Engineering & Technology, Karachi	DXX	2Mbps	2Mbps
54	University of Balochistan, Quetta	SDH	2Mbps	2Mbps
55	University of Karachi, Karachi	SDH	4Mbps	4Mbps
56	University of Sindh, Jamshoro	SDH	4Mbps	4Mbps