Master Thesis
Computer science
Thesis no: MSc-2007:16
June 2007

Meta-Model of Resilient Information System

Authors:
Adnan Ahmed
Syed Shahram Hussain

Department of
Interaction and System Design
School of Engineering
Blekinge Institute of Technology
Box 520
SE – 372 25 Ronneby
Sweden

This thesis is submitted to the Department of Interaction and System Design, School of Engineering at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Computer Science. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:
Authors:

Adnan Ahmed
E-mail: mianadnanahmed@gmail.com

Syed Shahram Hussain
E-mail: shahram_hussain@hotmail.com

University advisor:

Prof. Rune Gustavsson
Department of
Interaction and System Design
School of Engineering

Department of
Interaction and System Design
School of Engineering
Blekinge Institute of Technology
Box 520
SE – 372 25 Ronneby
Sweden

Internet : www.bth.se/tek
Phone : +46 457 38 50 00
Fax : + 46 457 271 25

ABSTRACT

The role of information systems has become very important in today's world. It is not only the business organizations who use information systems but the governments also posses' very critical information systems. The need is to make information systems available at all times under any situation. Information systems must have the capabilities to resist against the dangers to its services, performance & existence, and recover to its normal working state with the available resources in catastrophic situations. The information systems with such a capability can be called resilient information systems. This thesis is written to define resilient information systems, suggest its meta-model and to explain how existing technologies can be utilized for the development of resilient information system.

Keywords: Information Systems, Resilient Information Systems, Protection, Detection, Restoration.

ACKNOWLEDGEMENT

Thanks to Al-Mighty Allah and blessings to prophet Muhammad (Peace be upon him)

We would like to express gratitude to our supervisor, Prof. Dr. Rune Gustavsson, at Blekinge Institute of Technology for giving us the opportunity to work with him and for all his efforts, patience & encouragement for the successful completion of this thesis.

We would also like to appreciate the faculty at School of Engineering for the counsel and advice given to us during our studies here. We would like to thank all our friends in general and especially Sulman Mehmood, Musharrif Hussain, Sulman Majeed, Saad-bin-Saleem, Aamir Sohail, Ghazanfar Ali and at last but not least Waqas Siddique, for their help and support. We want you to know that we appreciate you and God bless you all.

Syed Shahram Hussain
Adnan Ahmed

I also like to say thanks to my parents and siblings (Imran, Saghir and Fatima) for their appreciation & support throughout my life. I would also like to say thanks to Prof. Muhammad Jamil for his endless support. A special thanks to my precious wife who has stood by my side all this time while making sure that the research will be a success.

Syed Shahram Hussain

I want to say thanks to my parents for teaching me my first words and for their support & love. I also want to thank my brothers (Irfan, Zeeshan), wife, in laws and my relatives for their appreciation and their support.

Adnan Ahmed

We dedicate this thesis to our parents.

CONTENTS

List of Figures

1 INTRODUCTION

This chapter provides the information about motivation, related work, aims & objectives, main contribution, research questions, research methodology and guidelines for reader.

1.1 Motivation for Writing Thesis

Information systems have become very important for the organizations and the individuals daily activities. The use of information systems have increased in the past of few years with the wide spread internet usage, rapid growth in computing power and the related integrating technologies. Organizations use information systems to perform their business operations (management, accounts, operations, etc) and individuals benefit themselves by using the services provided by these information systems (online shopping, online reservations, online banking, etc). At the same time these information systems are facing risks such as malicious programs, hacking, inadequate security policies, denial of service attacks, physical accidents, malfunction and out-dated system or software [1].

Due to this high demand and importance of information systems the need is to make information systems which can keep themselves intact and continue to provide services in catastrophic situations.

Making the information systems resilient can help in catering these obstacles. Resilience is an eligibility attribute to recover from effect or to resist some shock, or disturbance. Resilient information systems are such systems which come back in their original state after facing any catastrophic situation [2]. This attribute allows them to respond to sudden, unanticipated demands for performance and then return to their normal operating condition quickly with a minimum decrement in their performance [2]. The purpose of this dissertation is to find out different factors which can contribute to develop meta-model of information system which has resiliency in nature.

1.2 Related Work

The resilient information system is an emerging concept. In Europe the research has started in resilient systems and related area of study. European Union (EU) has assigned 9.1 billion Euro to fund Information and Communication Technologies (ICT) for the improvement of EU information and communication infrastructure [42]. The ICT has assigned 77 million Euro for the development of secure, dependable and trusted infrastructure [43].

In UK Central Sponsor for Information Assurance (CSIA) is working for the government to help maintain secure, reliable and resilient national information infrastructure. This research is being conducted in collaboration with their national and international partners in both public and private sector [1].

There is also a research being carried out in the design and maintenance of robust recoverable networks and information storage systems in Kentucky Center for Resilient Information Systems (CRIS). The CRIS is working in participation with University of Kentucky & Louisville, E-Cavern, IBM and Cisco and it is sponsored by US Department of Treasury [45].

Another research started in 2001 at IBM with the name Autonomic Computing [37], [44]. The idea behind this research is based on autonomic nervous system of human body. The main theme of autonomic computing is self management which can achieve by self protection, self healing, self configuring and self optimization [37]. The IBM is also researching on network resilience [7] and business resilience by having resilient infrastructure.

Sun has developed a new architecture for building and deploying systems and services capable of performing predictive self healing. Self healing is an important characteristic of a Resilient Information System Architecture to maximize the availability by implementing self healing at both software and hardware levels. Solaris 10 operating system is available with basic self healing features [48].

1.3 Aims and Objectives

The aim of this study is to define resilient information systems and to suggest how modern day information systems can be made resilient by using existing technology. The following are some objectives which will be addressed in this study:

- To find out the factors that arise the need to have resilient information system.
- To define resilient information system.
- To develop meta-model of resilient information system.
- To suggest the development of resilient information system by using existing technologies.

1.4 Research Question

Information systems cannot be left vulnerable because in case of emergency they require unmanageable time and resources to recover. It is an important task to make information systems resilient to failures so that even under unforeseen circumstances or under stress/attacks, information systems will behave as normal as possible and at the same time try to regain their previous stable state after heeling themselves [3]. The ultimate goal of such a mechanism is to develop resilient information systems capable of performing self assessment. By doing so, this knowledge can contribute in specifying new era of information systems.

The study is focused on the following research question:

- How to develop resilient information systems by using the characteristics of self protection, self configuration, self healing and self optimization?

The research question mentioned above is important enough to support a study because information systems are getting more and more endangered by different types of attack and there is a great need to make these systems resilient enough to bear or pass through such attacks.

1.5 Relationship between Research Questions and Objectives

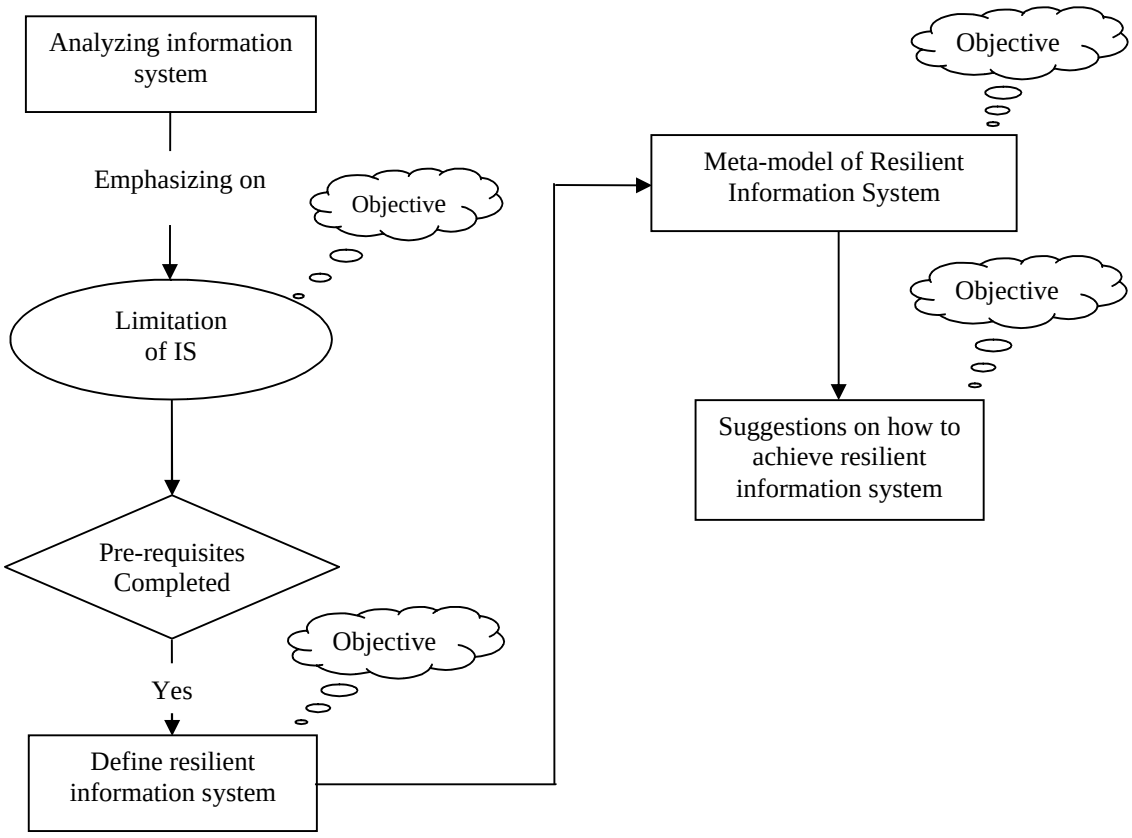

Figure 1.1: Relationship between objective and research questions

The preliminary stage of this study is to discuss information systems to find out their limitations. The knowledge about the limitations fulfills the pre-requisites for the study and also helps to achieve the objective of defining and developing a meta-model of resilient information system. The next stage is to gather the knowledge of existing technologies which are best suited for proposed meta-model of resilient information system. By doing so methods of monitoring, protection, detection and restoration are found. All of the above stages of the study help to accomplish the final objective of developing the resilient information system by using the existing technologies.

1.6 Research Methodology

This study employs qualitative methodology because the knowledge is developed by using constructivist claims [46]. The grounded theory strategy of qualitative method is used because it helps to produce general abstract theory of process [46]. This strategy helps to achieve the aim of the study and in answering the research question. That is transforming the information system to resilient information system. In strategy the data collection is performed in multiple stages of the study [46]. The data gathered in these multiple stages helps to define resilient information system and develop its meta-model.

Information systems and their limitation are discussed in perspective to highlight the importance of the need of resilient information systems. The best suitable platform to hold resilient information system and components of resilient information systems are discussed. To make information systems resilient components of resilience such as monitoring, protection, detection and restoration are used.

1.7 Main Contribution

The main contribution of this thesis is to find out the gap in existing information systems in terms of its ineligibility to cater the modern day needs. The need is to make the information systems resilient against any type of failure. To improve the information systems thus they can withstand and survive through catastrophic situations and make themselves eligible for the existing and future needs. The importance to have the resilience attribute in information systems is also highlighted. This is done by defining the resilient information system and then developed the meta-model of resilient information system. Then the further explanation of the meta-model is given by illustrating the major components of the resilient information systems. The explanation of the working and importance of these components are also discussed. In this thesis model of resilient information system is developed and based on that model the basic attributes of resilient information systems such as monitoring, protection, detection & restoration are discussed. It has also highlighted the necessity to transit from existing information systems to resilient information systems. This study is a good starting point and has opened up new interesting topics of research in the information systems and will contribute to make information systems more and more resilient.

1.8 Guidelines for The Reader

Chapter 2: This chapter covers the background of the thesis by addressing the information system and its limitations. The chapter starts with the definition of information system in section 2.1, following with addressing the architectural components of information system in section 2.2 and the different roles information systems plays in organizations is explained in section 2.3. The limitations of the information systems are highlighted in perspective to understand the need of resilient information systems in section 2.4.

Chapter 3: In this chapter resilient information system is discussed. The definition and explanation of resilient information system is addressed in section 3.1 and 3.2 consecutively. Maslow's theory of need is transformed to the information system need of resiliency in section 3.2. The meta-model of resilient information system is defined and explained by describing the functionality of its monitoring, protection, detection, restoration and knowledge-base in section 3.3.

Chapter 4: This chapter explains the proposed solution for development of resilient information. The infrastructure of resilient information system is addressed in section 4.1. Further more monitoring component is addressed in section 4.2. The protection component is addressed in section 4.3. Detection is addressed in section 4.4. Protection is discussed in section 4.5 and knowledge base is discussed in section 4.6. From section 4.2 to 4.6 the component wise solution of resilient information system is proposed by keeping in mind the infrastructure discussed in section 4.1.

Chapter 5: The main theme of this chapter to discuss the proposed solution.

Chapter 6: In this chapter future work of the research is explained.

2 BACKGROUND

The main theme of this chapter is to highlight the fact there is a great need to make information systems resilient.

2.1 Information Systems

> *"Information system is a system whether automated or manual, that comprises the entire infrastructure, organization, people, machines, and/or methods organized to collect, process, transmit , and disseminate data that represent user information* [5]*."*

Organizations and individuals at all levels and path's of life rely on information systems to manage their operations, compete in the marketplace, supply services, and also add a lot more to a persons life [40]. Corporations depend on computerized information systems to process their financial accounts and manage human resources, government organizations rely on information systems to provide basic services to its citizens and individuals use information systems to study, shop, bank, and invest [40]. Information systems have put forth a yawning influence over society by enabling more diverse human activities. These systems have accelerated the pace of daily activities, affected the structure of organizations, changed the way products are bought and sold, and influenced the nature of work. "Information and knowledge have become vital economic resources [40]".

Organizations are crossing all geographical and national boundaries and information systems play a very critical role in the race to expansion across the globe. These organizations use well designed information system architecture that supports modifications as new business opportunity arise [40].

2.2 Main Components of Information Systems Architecture

The main components of information system architecture are [41]:

- Equipment (computer hardware and software, databases, communication systems).
- People (human resources).
- The data/information.

Electronic information systems require computers to work on. Organizations depending on their need and the type of information system they have; use computers ranging from small size widely deployed personal computers (PCs) to large scale mainframes and supercomputers with supporting peripheral devices. To enable the hardware work there is a need to have software written for that particular hardware and also to provide basis for the applications including related to information systems. Information systems mainly focus on storing data in a systematic way and to do so databases are used. A database is a collection of related records organized in a manner that they can be retrieved efficiently. Communication systems are used to connect computer systems to transmit and share information. Internet is a

network of networks and is often used as communication network for information systems [40].

Humans are important component of information systems. Highly qualified personals develop, maintain and provide technical support to the information systems. Then there are people who act as the most important players the end beneficiaries of information systems called the users. Procedures for using information systems are also part of information systems [40].

The other part of information system is the data which is to be processed to make it available as information.

2.3 Roles of Information Systems

Organizations use Information Systems in three different roles [40]:

- As to support operations.
- As to support knowledge work.
- As to support management in organizations.

In operational role the information systems are transaction processing systems through which products are designed, marketed, produced, and delivered. These information systems build up information in databases that form the basis for higher-level systems [40]. Transaction processing information systems help to integrate supply chains of the organization by integrating the value chains of firms responsible for designing, marketing, producing and delivering the goods & services. Many transaction processing systems provide electronic commerce over the internet. Online shopping and E-banking are examples of such transactions processing systems [40].

A big percentage of work in information society relates to manoeuvre conceptual information and knowledge, other than directly processing, manufacturing, or delivering tangible materials and such a work is called knowledge work. The examples of information systems as working in a role of knowledge work are professional support systems, office information systems and knowledge management systems [40].

Information systems in another important role are as management support systems. In this role information systems are intended to support the management of an organization. Such information systems support all levels of management, from those in charge of short-term schedules and budgets for small work groups to those concerned with long-term plans and budgets for the entire organization. The management reporting system of management support systems present routines, detailed, and large information reports specific to each manager's area of responsibility [40].

2.4 Limitations of Information Systems

Yet, along with all these opportunities, information systems have exposed new threats [40].
Gradually the complexities in information systems are increasing. These complexities can be observed at all components of information systems. The need to have faster and faster computers with large storage capacity, well designed secure software and people with adequate knowledge have become the basic requirement for any information system. Computer networks are also growing because of the increasing trend in the use of

information systems in every aspects of our life [10]. People are becoming more and more dependent on computers, computer networks and information systems for the following reasons.

- The increasing computer power tends to make his life easy and simple
- The use of internet has decreased the cost of communication between two parties.

The organizations which are playing important role in world's economy are big customers of information technologies and generally have large information systems. Now a day business transactions depends upon these technologies. Any disaster natural are man made which may stop working of the infrastructure of information technologies resulting in information system failure can end up costing too much to these organizations and indirectly to the humans themselves.

In 1989, 50% of the business operations stopped in Sans Francisco Bay due to the earthquake, in result, businesses was shut down permanently and also the rate of unemployment increased [11]. According to the statistics provided in information security breaches survey held in UK in the year 2006, a large number of organizations have faced some kind of security breach, accidental systems failure, data corruption and premeditated & malicious incidents. This survey was conducted to find out the number of incidents occurred in last few years relating to security breaches, accidental failures, malicious activity recorded and their impact on the businesses. The following figure depicts the statistics of this survey.

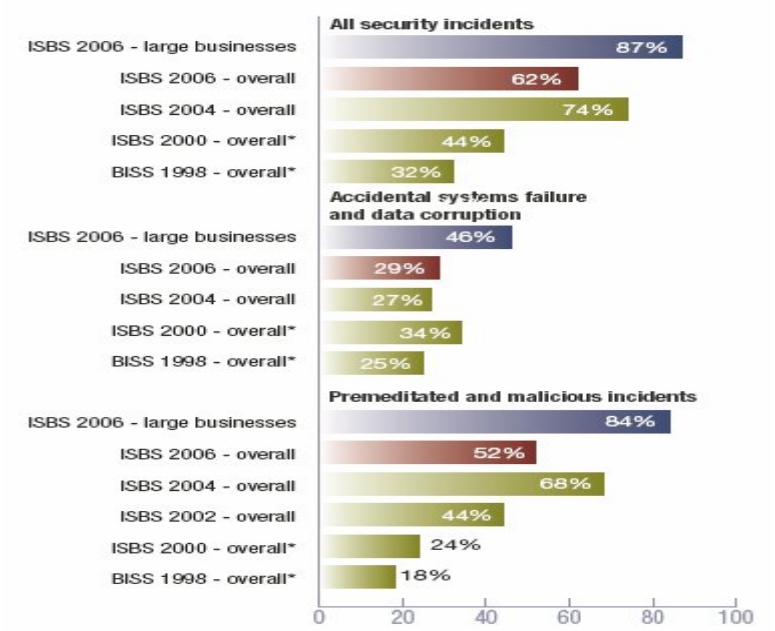

Figure 2.1: UK security incident which happened in 2006 [12]

The figure 2.1 clearly indicates the overall increase in number of security incidents in recent years. The figure 2.1 shows over all increase in premeditated and malicious incidents, accidental systems failure/data corruption and other security incidents happened over the years from 1998 to 2006. The number of premeditated and malicious incidents increased from 18% to 52% from 1998 to 2006. However statistics collected in 2006 only for large businesses are shocking as 84% of them got affected in one way or another. Accidental systems failure and data corruption increased from 25% in 1998 to 29% in 2006 but for large businesses again its alarming 46% for the year 2006 only. Obvious increase in all security incidents increased from 32% in 1998 to 62% in 2006 as a whole and again frightening 87%

for large businesses for the year 2006 only. A breach to large businesses costs on average £90,000 for the worst incident. The stats are not that much different in other parts of the world. Security incidents and systems failure costs businesses in terms of huge amounts. The following figure shows the amount spent and the losses.

	ISBS 2006 - overall	ISBS 2006 - large businesses
Business disruption	£6,000 - £12,000 *over 1-2 days*	£50,000 - £100,000 *over 1-2 days*
Time spent responding to incident	£600 - £1,200 *2-4 man-days*	£1,750 - £3,500 *5-10 man-days*
Direct cash spent responding to incident	£1,000 - £2,000	£5,000 - £10,000
Direct financial loss (e.g. loss of assets, fines etc.)	£500 - £1,000	£3,500 - £5,000
Damage to reputation	£100 - £400	£5,000 - £10,000
Total cost of worst incident on average	**£8,000 - £17,000**	**£65,000 - £130,000**

Figure 2.2: Due to security loopholes what companies faced in terms of cost during last year in UK [12]

The worst security incidents that caused the business disruption of a UK company costs roughly between (£8,000 to £17,000) during the year 2006. Large businesses have more security incidents and their breaches tend to be more expensive costing roughly between (£65,000 to £130,000) during the year 2006. Overall, the cost of security breaches to UK large businesses is up by roughly 50% since two years ago, and is of the order of ten billion pounds per annum [12]. Another important point to note is that IT infrastructure has a direct and indirect impact on the business operations of the organizations and it can be measured in terms of cost which organizations have to bear during the repair time in case of IT failure and the cost for ensuring the availability [7].

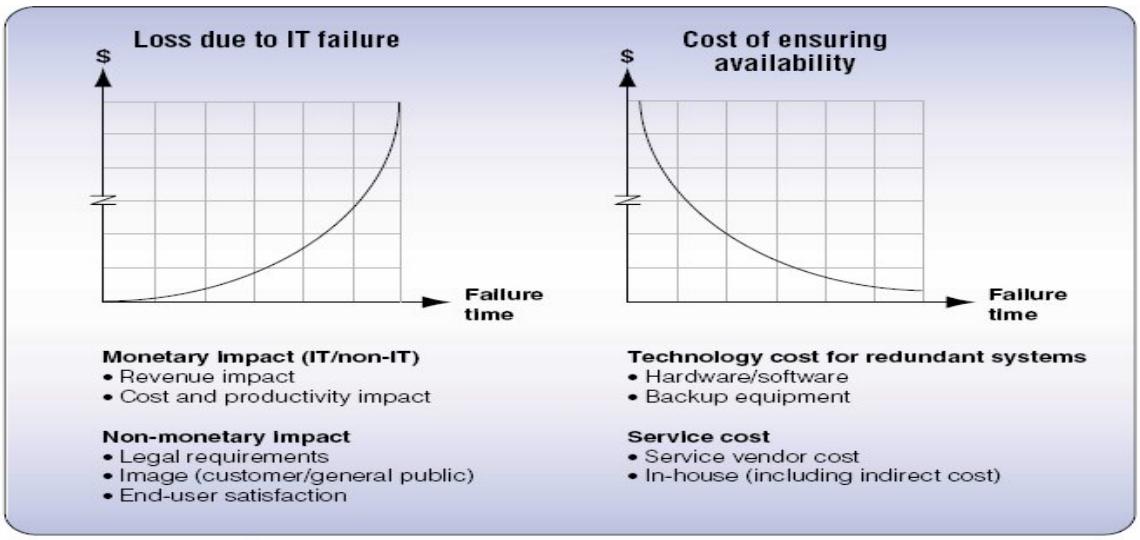

Figure 2.3: Weighing the cost [7]

According to above figure the longer the systems stay in state of failure, the bigger is the loss. In the figure one graph shows the loss due to IT failure increases as the failure time increases and the other graph shows the failure time decreases with the increase in cost spent for ensuring the availability.

It is not the business organizations who use information systems but the governments also posses very critical information systems and in general information systems worldwide are under constant threat of attack and the global reach of the Internet has magnifies this problem. New threats, more treacherous than the hackings of the past are arising. These new threats also include the threats driven by political motivations. Hacktivism, cyber terrorism and cyber warfare provide techniques to attack enemies across the world without ever having to cross a national border, or even leave home. Governments and private industry must understand the terrorization they face and get themselves ready to combat them [49].

By keeping this discussion in mind it can be said that there is a need to make information systems resilient, which can withstand to work and provide basic services resiliently under all circumstances. The ideal goal to have resilient infrastructure for information systems is to achieve zero failure time, however in reality it is difficult to achieve. Nonetheless, in practical resilient infrastructure for information systems can make a difference in an effort to reduce the chances of failure [7].

The figure below shows the transformation of an ordinary information system into a resilient information system.

Figure 2.4: Transformation impact [7]

When an ordinary system transforms into the resilient system, it is inclined to have a good impact on three areas such as people, organization and processes [7].

3 RESILIENT INFORMATION SYSTEM

This chapter defines resilient information system and the meta-model of Resilient Information System (RIS) is discussed.

3.1 Definition

> *"Resilient information system is such an intelligent system which takes pre-cautionary measures against attacks and if it is attacked and impaired then finds suitable ways to keep providing its functionality and restore itself to normal state with the available resources, it also keeps the information for this attack to refrain itself from getting into similar situation in future."*

3.2 Explanation of RIS

As it is evident from the above definition that the RIS must have three properties, first is the protection, second is the detection and third is the restoration. It can be resembled like any living being of the world. It must have an immune system working as it is in humans and also intelligent enough to find ways of its survival. The immune system of the human body provides protection against sickness/disease, and if unfortunately the body gets it, it tries to detect & heal it as soon as possible with minimal long term disabilities. Thus RIS must have the ability to protect itself against all type of system failure attacks. In case, the system unfortunately gets affected by an attack then detect the compromised resources and then it must have the ability to survive through it with as little damage as possible, and perform restoration itself by using remaining resources.

Humans learn by their experience and try to avoid these situations which can harm them by using the information they gathered from previous catastrophic situation. RIS must also have this ability to learn from the past experience and identify similar attacks which have harmed it in the past or identify and protect from those attack which may affect the RIS.

After having this little understanding of the concept to have RIS, few questions arise and one of them is: what is the basic structure of resilient information system? To have the answer of this basic question there is a need to look around in civilized societies, big cities and the whole world. Every manageable thing is governed by some organization or some group of people who are taking care of all these things. For example traffic system in all countries have some basic common rules. Every vehicle which is on the road must follow these rules to run the system smoothly. Any body found not following these rules is warned or penalized by the traffic department. Humans live in communities and communities have their own rules and regulations. People living in these communities follow these rules and regulations of that particular community.

Likewise RIS must also be organized in such a way that rules and regulations can be applied to it. This arise the concept of community base resilient information systems. It is easy to handle traceability and accountability matters in a community. The traceability in resilient information system is to identify the vulnerable area of the system and accountability is to fix responsibility. When these things are handled in a community then it does not matter how

big the community is because every member of that community will follow the same rules, technologies and structure of the RIS.

The motivation to have the RIS can be found through the psychological theories such as Maslow's hierarchy of need which is proposed by an American psychologist in 1943 [9].The following figure 3.1 depicts this theory.

Self-actualisation	The craving of beauty, creative activity and order. The desire to know and understand. The need to feel self-fulfilled.
Esteem	The need for high evolution of self and recognition from other, leading to self confidence and adequacy.
Belonging	The need to give and receive affection and relate to other people.
Safety	The need to ensure freedom from danger and establishing stability.
Physiological	The need to satisfy basic requirement before progressing further such as hunger and thirst.

Figure 3.1: Maslow's theory of needs [9]

Maslow's hierarchy of basic human needs define several levels of needs based directly on top of each other. A person's most basic needs are those required to sustain life such as air, water and sleep. Once the basic needs are met, the person's attention will turn to its safety and security. The examples of such needs are living in a safe area, medical insurance, job security and financial reserves. Once a person has met the psychological and safety needs it will start interacting with other people to give and receive affection. At this level the sense of belonging will come to the person. At this level the needs of a person will be concerned about its esteem such as self-respect, achievement, attention, recognition and reputation. The top layer in Maslow's theory of needs is self-actualization. It is a quest of reaching one's full potential. The needs of this level are never fully satisfied as the person's needs at this level grow all the time. The examples of needs of this level are truth, justice, wisdom, meaning and etc [46].

This theory of need can be related to RIS as shown in figure 3.2.

Maslow's Theory of Need	**Resilient Information System**
Self-actualisation	Awareness of its components
Esteem	Knowledge of its power ,system Identity
Belonging	Co-operation with other sub system
Safety	Self-protection
Physiological	Self-Configuration Self-Healing Self-Optimising

Figure 3.2: Correlation of Maslow's theory and RIS [9]

Maslow's theory of need can be compared to resilient information systems by its characteristics such as self protection, self healing, self configuration and self optimization. Maintaining basic system operations through self healing and self configuration in RIS to

achieve optimization corresponds to the physiological needs. Maslow's hierarchy gives a guideline for resilient information systems in terms of priorities and the fulfillment of objectives according to the theory of needs. Self protection corresponds to safety needs of the system in accordance with the theory of needs [9]. Communication with internal components of RIS as well as communication with external networked resources and data sources corresponds to the need of belonging. In RIS the decision to perform or not to perform an action depending on the situation corresponds to esteem and self actualization. The awareness of the new situation also corresponds to esteem and self actualization of needs theory [9].

The important characteristics of self protection, self healing, self configuration and self optimization of RIS according to the Maslow's theory will be focused in this thesis.

Self Protection is a term used for the protection of system against all type of attacks. It is like a boundary which is used to protect against the enemy.

Self Healing is a term used for automating the process of identifying and fixing frequently accruing problems.

Self Configuration is a term used for an attribute that can learn to modify its present configuration in reaction to the current workload. It is like traffic system which has traffic signals to formulate the flow of traffic and then if an accident happens at a particular route the traffic department configures and formulate another route so that the flow of traffic would not disturb.

Self Optimization is a term used for analyzing the situation, determining the goals and adjustment of the behavior.

The IBM is one of the companies providing the solution about resilience infrastructure. It has also suggested a pro-active technique for planning toward resilient computing environments. The components involved in planning can be divided into six resiliency layers as is shown in the figure 3.3.

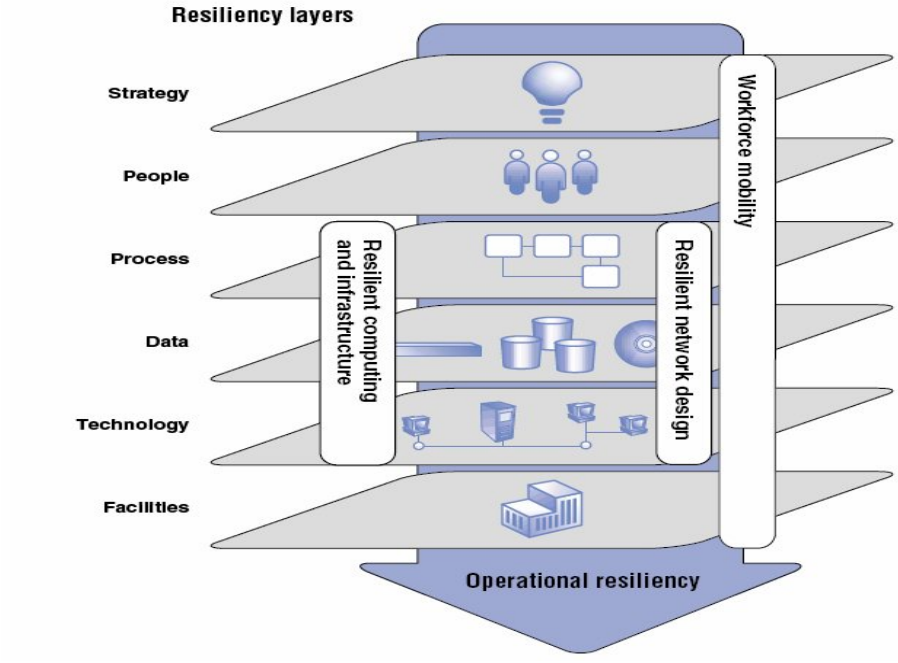

Figure 3.3: Planning Operational Resilience [7]

Moving towards operational resiliency involves determining the strategy, making preparations by training people, devise new smart processes, improve data flows, and develop fail proof technology. The basic aim to have resiliency is to increase the ability of systems to support changing or unexpected conditions. It involves providing reliability, security, sustainability. Resilient computing infrastructures include making strategies, involving people, processes and technology altogether in way that they can sustain under changing and unexpected conditions [7].

To make such a resilient computing infrastructure work there is a need to have a resilient network as is discussed in the figure 3.4 given below.

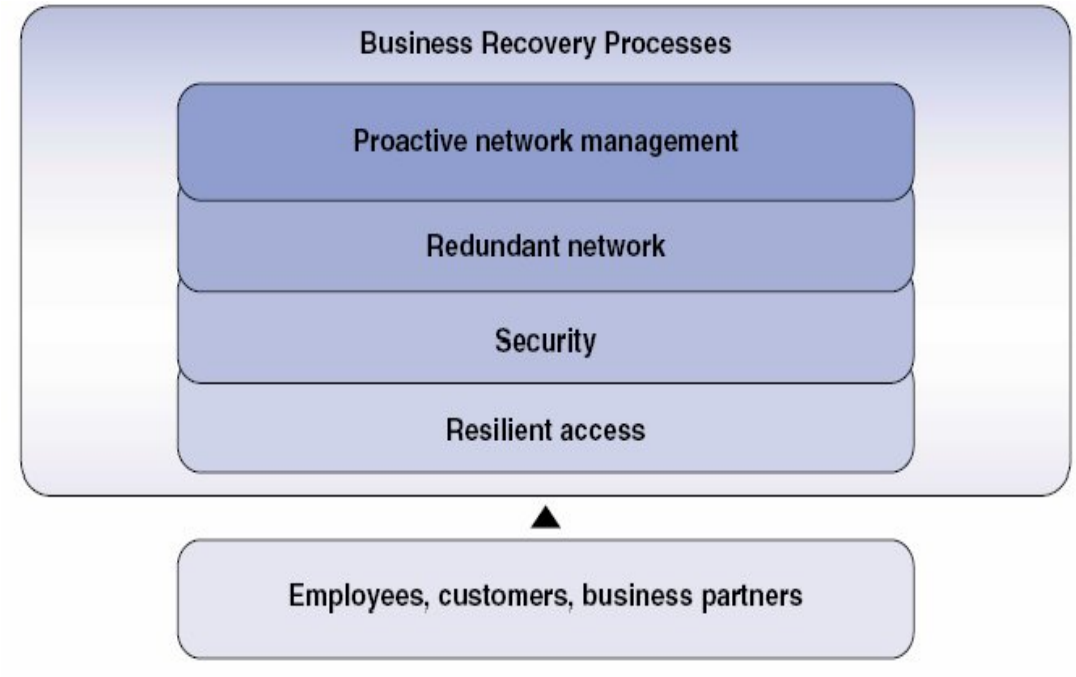

Figure 3.4: Resilient network design [7]

The above diagram depicts the major elements of resilient network design. A resilient network is an important part of the resilient infrastructure. Resilient computing infrastructure is accessed through resilient networks. To make a network resilient there must be a resilient access mechanism and then security checks must be performed. A redundant network will look after the failed nodes and optimizing the node usage with the help of proactive network management.

3.3 Meta-Model of RIS

The figure 3.6 is a proposed model of RIS. The model is used to portray the actual meaning of the definition of RIS. It also explains the essential components of a RIS and the relation between these components in order to achieve the characteristics of RIS.

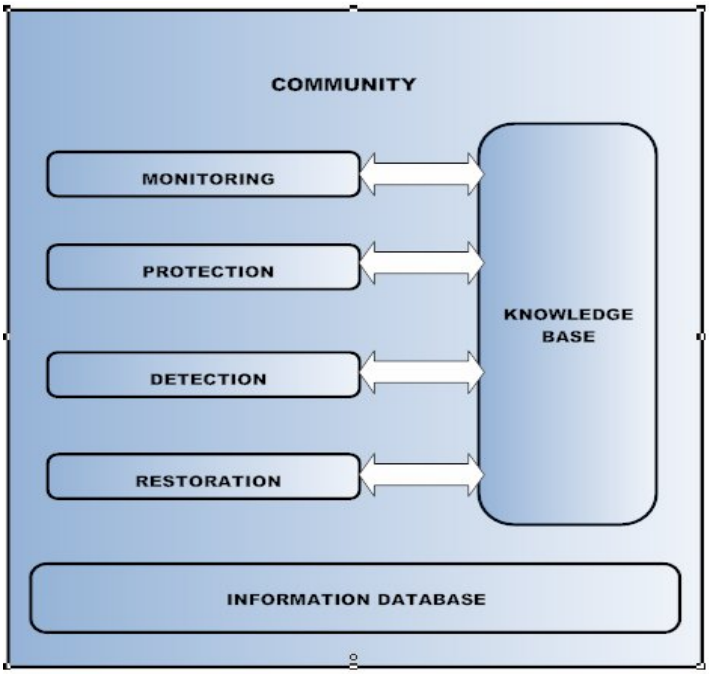

Figure 3.5: Meta-model of RIS

The above model is explained further by discussing its components.

3.3.1 Community

Community is a group of people sharing the concepts and having the same interests. In the above model it means that a RIS must work in community so they will follow some standards to achieve resiliency. The benefit to have a community is that every community has its predefined rules and regulations which are mandatory to be followed by all the members of that particular community.

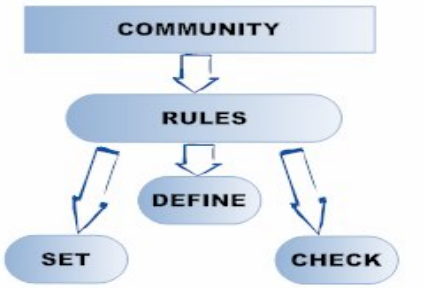

Figure 3.6: Community of RIS

RIS community will define, amend, checks rules for the implementation, operations and up-gradation of RIS.

3.3.2 Monitoring

Monitoring of the RIS is an essential component; it performs monitoring on operational resources, data flow, components to help protection, detection and restoration components to work. Without monitoring there can be no concept of RIS. The protection, detection and restoration components are heavily dependent on data gathered by monitoring component.

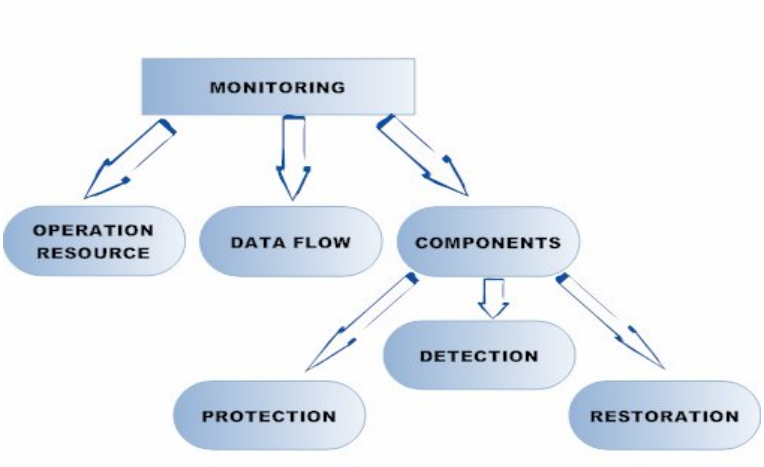

Figure 3.7: Monitoring Component

Monitoring is performed by collecting data from different sensors working at operational resources, dataflow and with components. The monitoring data will be stored in Knowledge base for other components to retrieve and perform the necessitated operations.

3.3.3 Protection

Protection is the key component of RIS. This plays both the role of defence and offence in RIS against all types of attack. The defense part is to protect from attacks and keep updating the knowledge base for new attacks. The offence part is that if RIS has been harmed by some attack in the past and the same attack occurs again; it will protect the RIS from it.

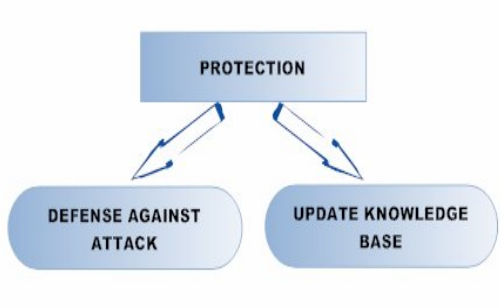

Figure 3.8: Protection Component

3.3.4 Detection

Detection detects the vulnerabilities and the weak points of the RIS. The processes of detection component are initiated by monitoring to find the causes of failure at some part. The data gathered by detection is stored in knowledgebase. This data is utilized by restoration component in order to perform restoration and optimization of RIS. To perform detection different detection algorithms depending upon the scenario will be used. The

vulnerabilities detected will update the vulnerability list in knowledgebase for future protection.

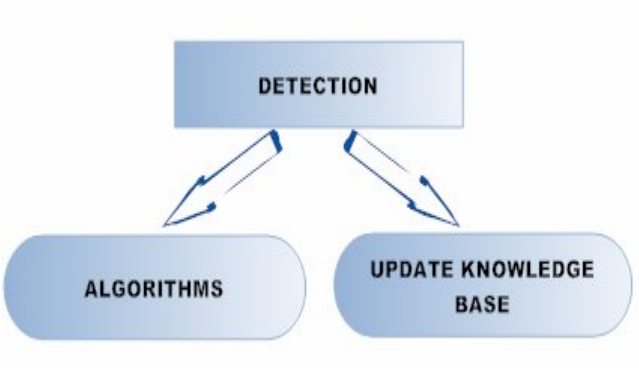

Figure 3.9: Detection Component

3.3.5 Restoration

The main responsibility of restoration component is to bring back the RIS to its normal state after catastrophic situation. The restoration component will perform healing on weakened parts of the RIS and empower them to perform their regular functions. The weak parts will be detected by the detection component with the help of the monitoring component. Self configuration will be used to reconfigure the components when workload is increased to achieve optimization of RIS. Optimization will also be used in case when full restoration is not achieved by the restoration. The restoration component has following properties as shown in figure 3.11.

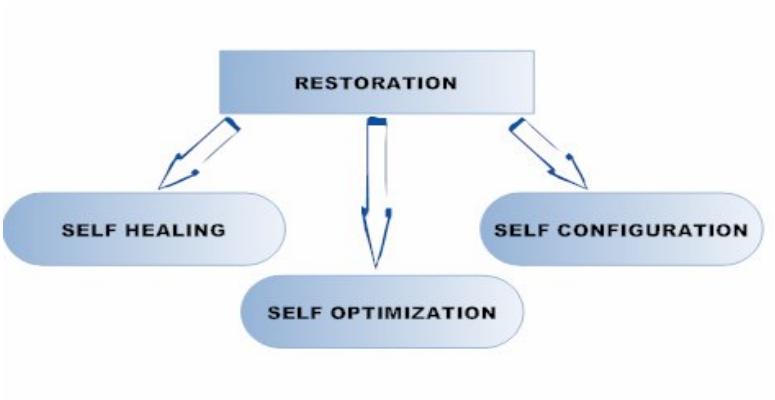

Figure 3.10: Restoration Component

3.3.6 Knowledge Base

The knowledge base component of the RIS is a data base that keeps the information such as monitoring data, restoration data, vulnerabilities list, and information about all types of attack and its precautionary measures. Monitoring data is collected by the sensors working with all the components. This data will be used by other components to perform their operations. Detection data is gathered during detection phase. This data includes the detected faults and this data will be used by restoration to perform restoration of RIS.

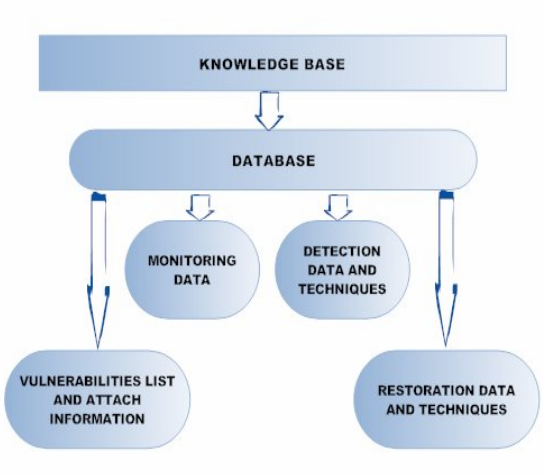

Figure 3.11: Knowledge base component

3.3.7 Information Database

Information database is the basic resource which is shared on RIS.

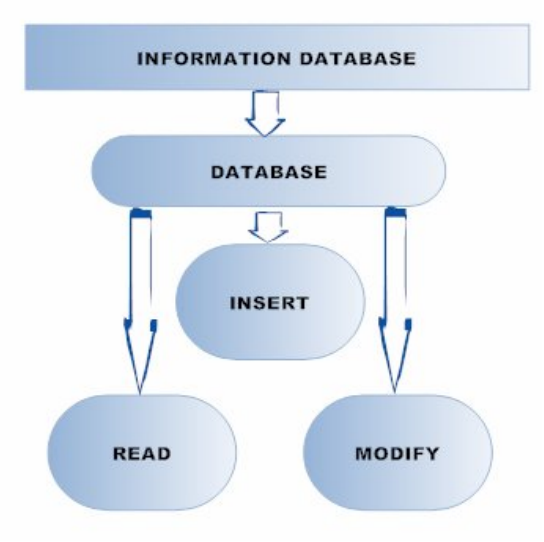

Figure 3.12: Information Database

4 PROPOSED SOLUTION

This chapter is written keeping in mind how to make a resilient information system using the existing technologies. Which infrastructure is best suitable for the implementation of RIS meta-model. Which are the technologies available for implementing protection, detection and restoration.

4.1 Infrastructure of RIS

The infrastructure of RIS is community based as it is depicted in figure 3.1. To understand it, let us look at what is the meaning of community. The word community means the group of people sharing the concepts and having the same interest. The question may arise what people have to do in RIS infrastructure. The answer is to have such a dream system we need group of people, software's and hardware's for the management and operations of the information system in a resilient way. To cater this there is a need to bring people & the resources at one platform and make a community. The benefit to have a community is that every community has its predefined rules and regulations which are mandatory to be followed by every person belonging to that particular community. Such rules and regulations help the community to cater the problems like accountability, traceability, extendibility and adoptability.

By keeping in mind the existing technologies, it seems obvious that grid-computing infrastructure is best suited infrastructure for the development of RIS. The grid computing is an emerging technology in the field of distributed computing. The term grid describe by Ian Foster of Argonne National Laboratory and Cal Kesselman of the University of Southern California ,first in workshop at Argonne National Laboratory in September 1997 and after in publication "The Grid: Blueprint for a New Computing infrastructure" in 1998[13]. The definition of grid computing is [14].

> *"Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations."*

According to IBM the grid computing is [15].

> *"Grid computing enables the virtualization of distributed computing and data resources such as processing, network bandwidth and storage capacity to create a single system image, granting users and applications seamless access to vast IT capabilities. Just as an Internet user views a unified instance of content via the Web, a grid user essentially sees a single, large virtual computer."*

According to Ian Foster following are the basic characteristics a gird infrastructure must possess.

1. The grid has the ability to integrate, to coordinate resources and users of different networks which have different policies and administration control [16].

2. When building the grid the protocols and interfaces are used is open and standard. The protocols and interfaces have the capabilities of catering the problems of authentication, authorization resource discovery and resource access [16].

3. The grid allows its components to be used in coordinated fashion to get the quality of services. Such as response time, throughput, availability and security [16].

The architecture grid is also described in layers; in which each layer performs its own defined functionality. The figure 4.1 depicts the layered architecture of the grid.

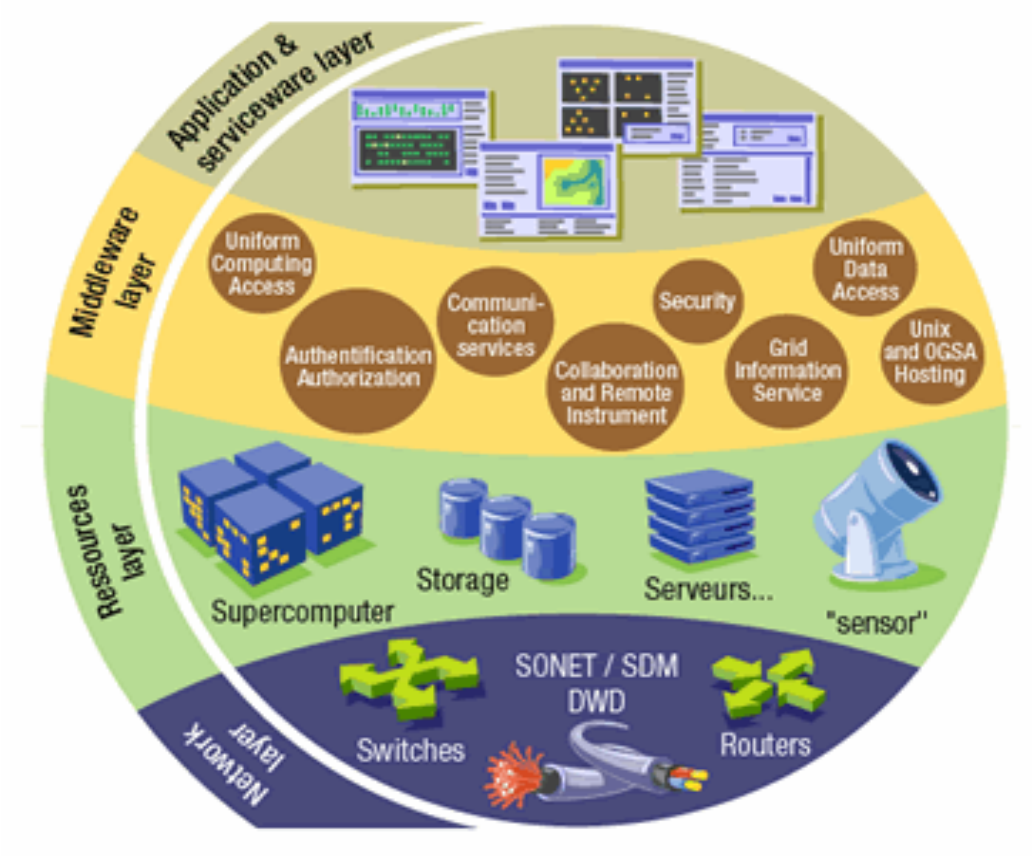

Figure 4.1: Layered Architecture of the Gird [17]

The lower layer is a network layer which deals layer of the grid architecture handles the problems related to part of the resources with all related issues of networks and responsible for connectivity between the resources in the grid [17]. Resource in the grid such as computers, electronic data, sensor, telescope and storage devices, which might be connected to the network [17].

The middle layer of the grid architecture is also called brain of the grid [17]. Its responsibility is to provide the tools for different components of the grid which helps these components to participate in the grid environment [17]. The top of the grid architecture is application and service ware layer; this is the layer which user interacts with because other layers of the gird architecture are invisible from the user [17]. The responsibilities of this layer is keep track about who is using these resources of the grid and who is providing these resources and charging the bill for these resources etc [17]. There is also another why

describing the layered architecture of the grid. The following figure 4.2 shows this architecture.

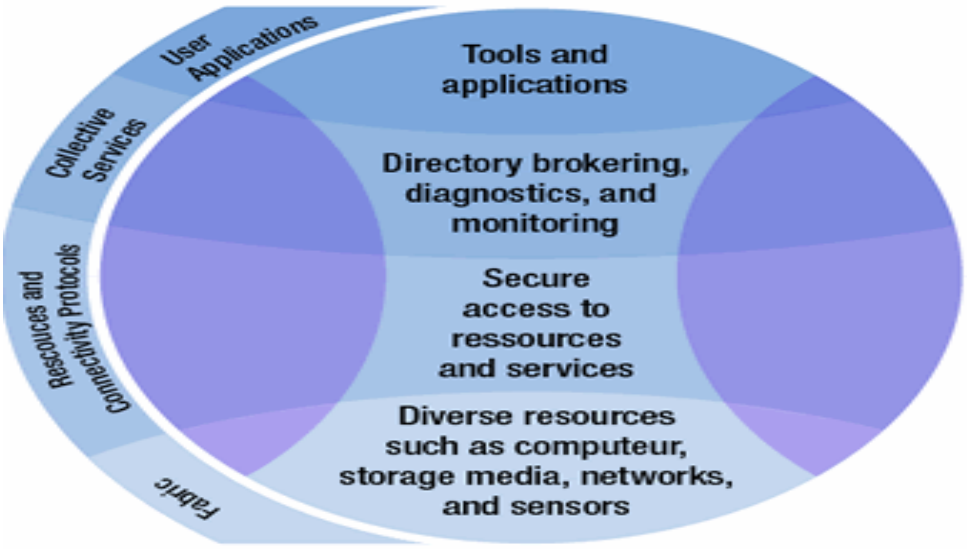

Figure 4.2: Layered Architecture of the grid [17]

The fabric layer of the architecture showed in figure 4.1 deals with the physical aspects of communication of network and computer [17]. The resource and connectivity protocol handles the issues of network transactions of the grid between the resources and different computer of the grid because the gird uses the Internet and there are so many other Internet services taking place at the same time for example email [17]. Therefore for a smoothly performing work on grid the communication protocol is needed which helps grid computers to understand those messages which are relevant to them [17].

The collective layer of the gird architecture showed in figure 4.2 also has protocols [17]. One of them is an information protocol that helps to gather the information of the structure and grid resources [17]. The other one is management protocol, which helps in consistent way to negotiate to access of the resources [17]. The application layer helps to handles the application of gird in better way [17]. The following figure 4.3 shows the relationship of the layered architecture and internet protocol.

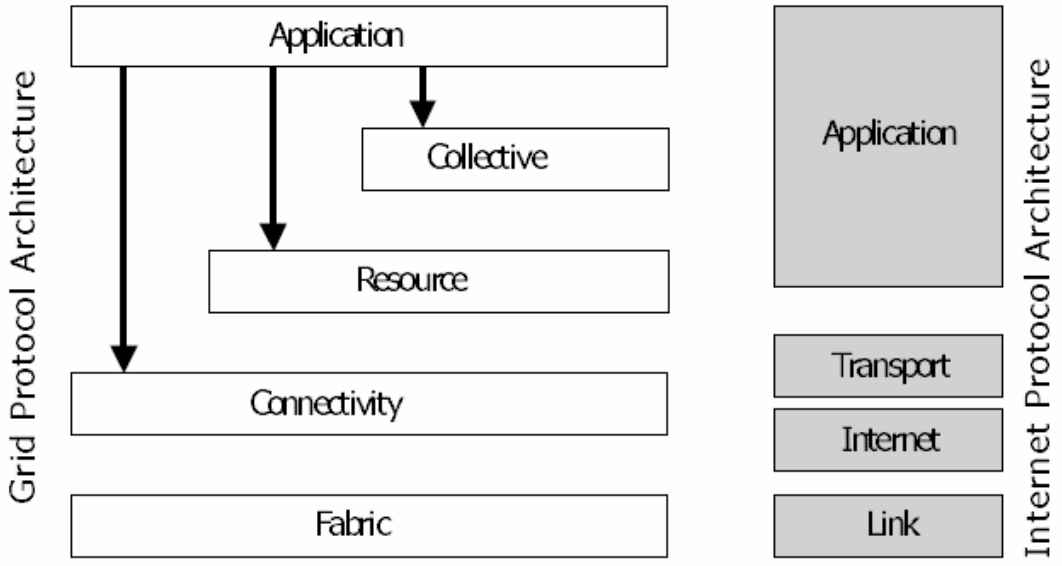

Figure 4.3: Relationship between layered architecture of the grid and internet Protocol [14]

Following are some benefits of using the grid which is addressed by IBM [18].

1. Use of underutilized resources

Now days the computer is everywhere and most of the computer resources are not utilized properly. The utilization of the desktop computers in big organization is 5 % and some of the organizations even servers are not properly utilized [18]. When these organization are going to join the grid then resources will be utilized properly.

2. CPU power utilization

The CPU power is needed where more computation has to be performed such as scientific research, motion picture animation, oil exploring , bio medical and financial modeling [18]. For these situation the gird computing is ideal and cheap. Only to do is develop such an application which uses an algorithm that divides the task in sub tasks and these sub tasks divide in such a way that they do not need to communicate with each other [18]. These sub tasks can execute on different nodes of the grid [18].

3. Virtual organization

The expandability of the grid is very easy due to virtual organization. The organization, institute and research laboratories can join the grid; however they are using different administration policies or might be using different operating systems by creating the virtual organization. The collaboration of virtual organizations develops the image of virtual computing which is providing various virtual resources [18]. The figure 4.4 depicts this concept.

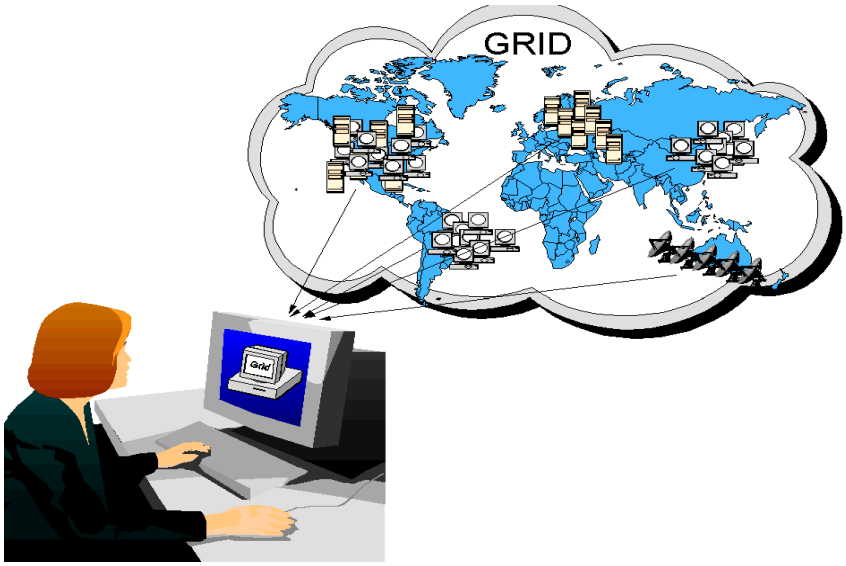

Figure 4.4: Collaboration of virtual organization [18]

4. Utilization of additional resources

The grid also provides access to the other resources utilization as well as traditional resource sharing (CPU power, storage) [18]. There are many other resources: some of them are special instruments and soft wares [18]. For example, the maximum band width of the grid can be used if the user wants to implement the data mining search engine [18].

5. Balancing the resources

The large amounts of the resources are distributed by the grid; these resources are provided by individual machines of the grid [18]. The grid facilitates applications by performing resource balancing on job scheduling of the application processes [18]. These applications must have the capability to run on the grid [18]. The resource is balancing perform by the grid in two ways [18].

- The unexpected processing power can be transfer to idle machine [18].
- The other scenario is that when gird resources are fully utilized then resource balancing is performed by giving the resources of the gird to high priority work and suspends lower priority until the higher priority work done [18].

6. Reliability

The reliability of the any system is a hot issue now days. The companies are spending so much money now days buying expensive hardware to increase the reliability of their system [18]. The grid infrastructure solution is inexpensive solution as compared to other solutions for reliability [18]. When any location of the grid goes down the grid management application resubmits the gobs of this location on another location of the grid [18]. In critical environment same jobs are running on different locations to handle the reliability issues[18]. The following figure 4.5 depicts this scenario

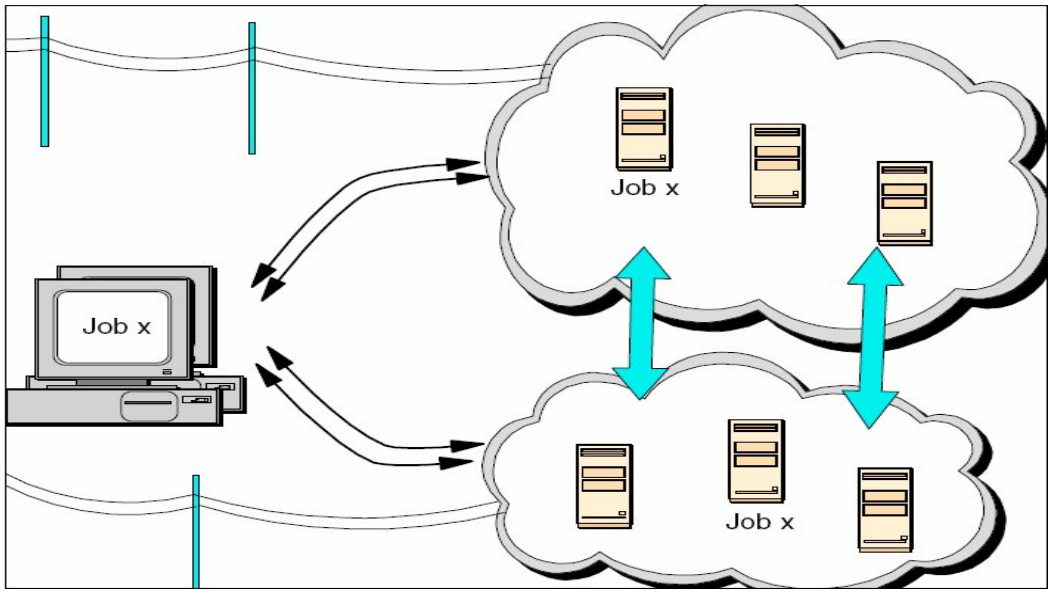

Figure 4.5 : In critical environment reliability structure of the grid [18]

7. Management

The sources become virtualized in grid computing due this management of the resources become easy [18]. It also helps to visualize the capacity and utilization of the resources in order to decrease the cost of the computing resources in IT infrastructure [18]. In different projects grid provides management priorities which has not happened in the past [18]. In past every project have their own computing resources which is not shared by any other project, due to this unutilized resources can not be used by other projects whose needs are growing due to unexpected event [18]. As the grid viewed as a virtual computer the resource

management issues can be handled easily and also help the administrator to changes the policies of the resource utilization of resources by the organizations [18].

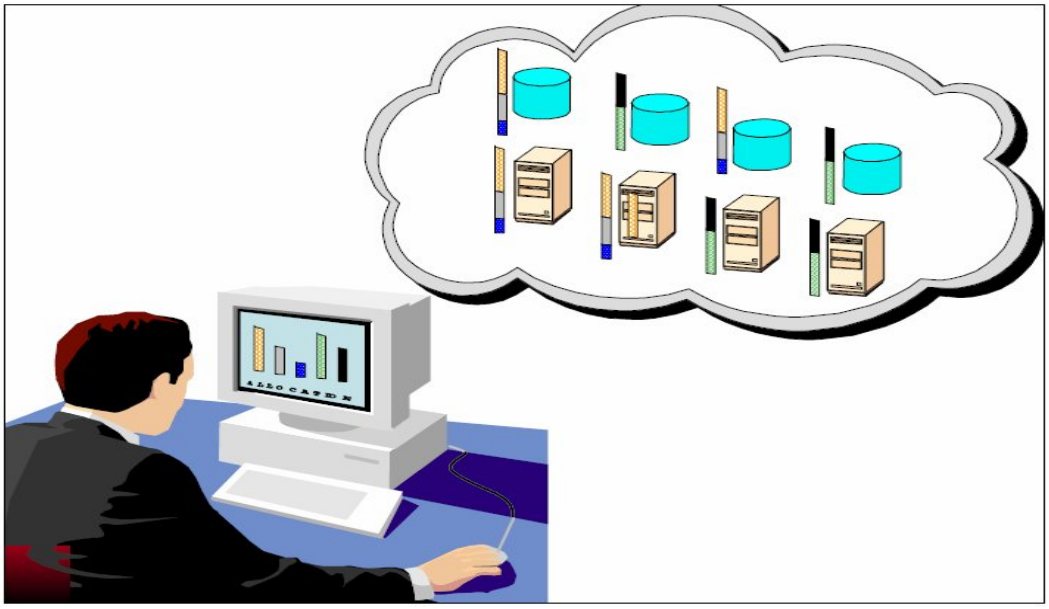

Figure 4.6: Resource management operations [18]

To development of the grid there is need of protocol or architecture. For this the Globus Toolkit (GT) [19] is open source community base architecture [20]. GT provides open source libraries for the development of the grid application [20]. It also handles the problems related to portability, resource management, security, data management, and communication and information discovery [20]. The most of the gird project uses the GT mechanism [20].

By the development of open grid services architecture (OGSA) [21, 22] new change is brought by in grid industry, it has concept and technology based upon web services [19]. It also increase the interoperability in the gird [22]. The capabilities that OGSA has are depicted in figure 4.5.

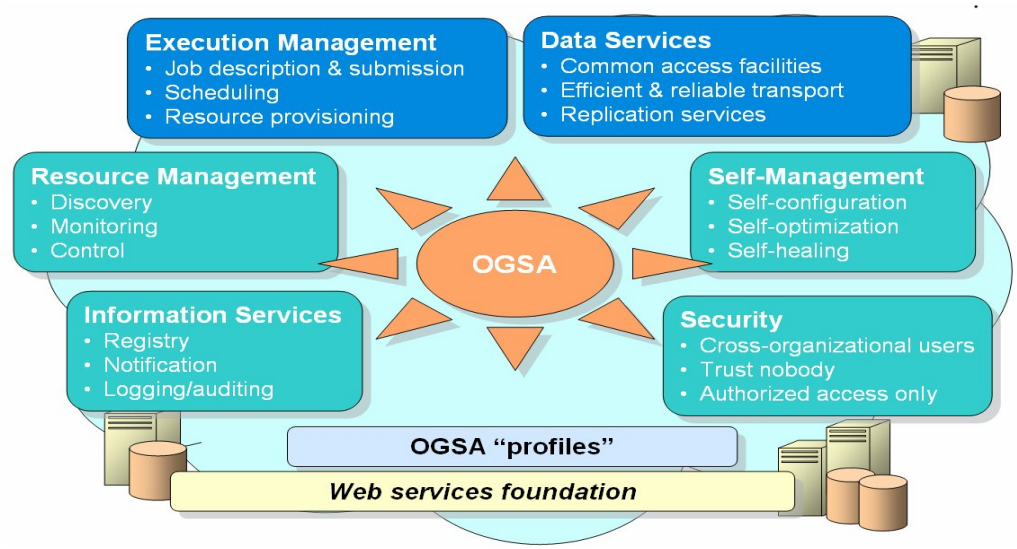

Figure 4.7: Capabilities of OGSA [23]

The OGSA capabilities are used in globus project for development of GT version 3 (GT3) [20]. The GT3 is the implementation of OGSA [20].

4.2 Monitoring component of RIS

The RIS is based on gird infrastructure due to this the grid monitoring techniques are used to performing the monitoring operation on RIS. Due to the virtual nature of the grid, the resources which are part of the gird join and leave dynamically [24]. The purpose of the monitoring is gathering and providing the status information of the grid [25]. The following are the benefits of monitoring [26].

- **Analysis and fault detection:** The data gather by the monitoring process helps to determine the status of the job execution such as job is executing fine, died or hung [26].

- **Monitoring job progress:** User which perform complex job on grid can easily get the status of their executing job from the monitoring data and this data is can also used in graphical tool to show job status [26].

- **Performance analysis:** The précised monitoring data help to determine bottleneck performance analysis [26].

- **Network aware application:** The application can also used monitoring data to tune itself, increase performance execution itself and configure itself [26].

- **Services of data replication:** Migrating the data to other location these services require the monitoring data [26].

- **Auditing:** The monitoring data also help in auditing the resource utilization and this auditing help in billing process [26].

- **Prediction and scheduling services:** To assigning the optimal resource to job the monitoring data can be used and it is also used in scheduling process [26].

- **Monitoring configuration :** To keep monitoring process run perfectly the monitoring keep updating its configuration by using monitoring data [26].

- **Restoration, detection & protection:** In restoration, detection and protection of RIS component also need monitoring to performing their process successfully.

This monitoring component keeps entering the monitoring data in RIS knowledge base. When other components need monitoring data, they get it from knowledge base. The monitoring system is consisting of following components [24].

- **Sensor:** In the monitoring process the sensor performs the actual monitoring [24]. Data that is gathered by the sensor is called event [24]. There are many different types of the sensor such as network, general, host, application and custom made [24].

- **Sensor management:** This component helps to control and configure the sensor remotely [24].

- **Information database:** This component stores the information about available monitoring event and sensor [24].

- **Data archiving:** This component is necessary for deep performance analysis, resource utilization statistics and accounting.

- **Presentation:** This component helps to view graphically representation of data.

The Global Grid Forum (GGF) purposed Grid Monitoring Architecture (GMA) for monitoring the gird [24].This GMA is composed of three components which are producer, directory service and consumer [24]. The figure 4.6 shows the components and their relations between them.

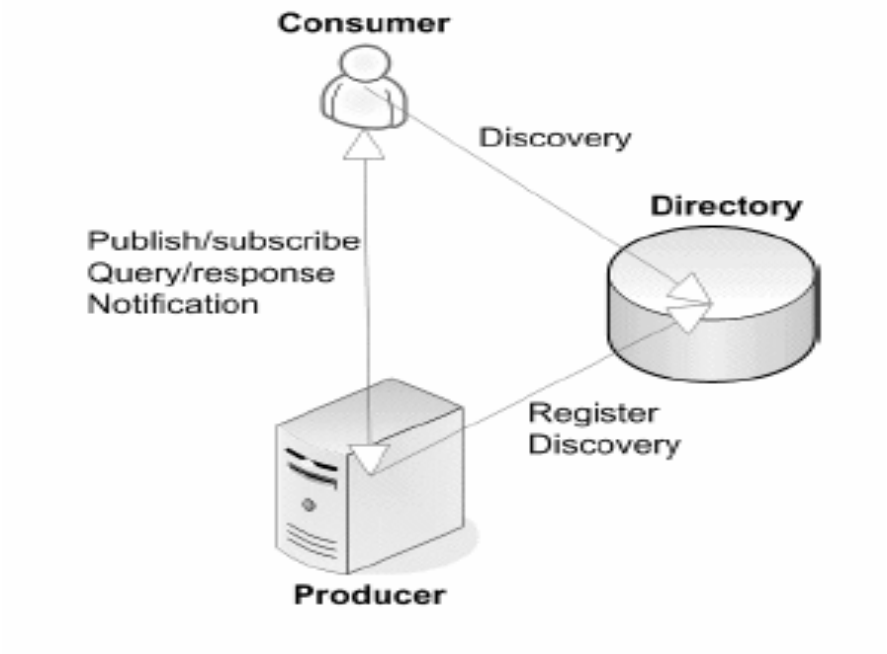

Figure 4.8 : GMA components and their relation [24]

The entity that demands the monitoring of specific resource is called the consumer service [24]. The entity may be grid, user, application and subsystem [24]. The service that represent publishes event and sensor is called producer [24]. Directory service in GMA is used for storing the information about consumers and producers [24]. Procedure used the directory service to store the information about monitored events and consumer used to register the events of its interest [24]. It is also used in discovery the interested events for consumer and producer [24].

There are some basic requirements of grid monitoring system which are described as following [24].

- **Sensor management**: In this monitoring system provides service of remotely configuration and controlling of the sensor [24].

- **Discovery and publish**: The sensor of the monitoring system must publish the information about resources and help to user in discovering the information about resources [24].

- **Non-intrusiveness**: During the designing of the monitoring system, it must be considered that monitoring system does not get affected by the measurement process [24].

- **Security**: The monitoring system must have the capability to provide the basic which are secure message flow, authorization and authentication [24].

- **Interoperability**: The monitoring system must cooperate with other grid middleware system and monitoring system [24].

- **Data archiving**: There must be repository for monitored data, which can be utilized in resource utilization statistics and application performance analysis [24].

- **Scalability**: The monitoring system has the capability to support large scale IT infrastructure or dynamic environment [24].

- **Extensibility**: The monitoring system is designed in such a way that is has the ability to extend basic functionalities of the system when it is needed [24].

Some of the monitoring systems are described as follows:

Relational – Grid Monitoring Architecture:

European Data Grid project developed Relational Grid Monitoring Architecture(R-GMA) as a gird information and monitoring system. It is based one GMA which is developed by GGF [27]. In R-GMA there in no need to know about the registry for those which are supplying the registry and getting information from R-GMA. The reason is that in R-GMA procedure and consumer handling the registry behind the scene [28]. To adopt OGSA framework the R-GMA migrating to web services [27].

Monitoring Agents in a Large Integrated Services Architecture:

Monitoring Agents in Large Integrated Services Architecture (MonALISA) is developed by Caltech and its partners over the four year with the help of U.S. CMS and computing program [29]. It provided the distributed service architecture which based on process and collect monitoring information. MonALISA is implemented in WSDL and in java using JINI technologies [30]. All aspects of monitoring is performed by MonALISA on complex system such as [31].

- Gathering the information for Computer node and cluster system

- Gathering the network information such as WLAN and LAB.

- MonALISA perform monitoring for job and the performance of application.

- The performance of end user systems.

In MonALISA framework the specialized mobile agent is used to improve the operation of large distributed system by supervising tasks of different applications or real time parameters and also perform global optimization task [30].

Monitoring and Discovery System:

Monitoring and Discovery system (MDS) which is part of the GT of Globus Alliance [19].The purpose of the MDS is to provide simple host monitoring sensors and information management [24]. Usually the Ganglia [32] and Network Weather Service (NWS) [33] are monitoring system which provides sensors to MDS. Up to now there are four versions of the MDS is released. MDS2 which is part of Globus version 2.x and based on LDAP. The MDS3 & MDS4 is part of Globus version 3 & 4 respectively and their implementation based on emerging gird architecture OGSA[24].

As the infrastructure of the RIS is grid and the grid is composed of the virtual organizations which are administrated by different companies. These companies have different polices for resource utilizations. Due to this availability assurance of the resources might be decreased and also impact on the grid performance. Problem of the commitment and assurance is handled with service level agreement (SLA) [34]. The SLA is contract between service providers and users who are using the resources. In SLA the users and grid service providers explicitly define the statements of expectation and obligations in business relationship between users and gird service providers[35].

When these SLA are used in grid then monitoring needs arises for SLA. The monitoring data which is gather from by performing SLA monitoring used for fault detection. It can also use for performing the checking on virtual organizations of the grid to assuring that these virtual organizations sharing the resources according to SLA or not. Doing back tracking of SLA using the monitoring data performs the fault detection. The protection component of RIS can also use monitoring data of SLA for authentication and authorization. Due to these reasons the monitoring system must have the capabilities of performing monitoring on SLA.

4.3 Protection component of RIS

The protection component is one of the most important components of RIS. It is like the countries borders which are protecting the countries from the illegal entry of persons, the forces of other countries and terrorist. The hundred percent protections is difficulty to achieve but as the protection percentage increases the reliability of the RIS is also increased. Powerful protection component also minimizes the other components such as detection and restoration. Protection component focus on security issues of the RIS. As the RIS infrastructure is based on grid. The following are the grid security challenges [36].

Integration Challenge:

It is impossible to give the single security solution that helps to cater all the grid security challenges .The reason is that it is impossible to replace existing technologies with new solution of the grid security immediately [36]. For example there are users account registries in each network domain in grid infrastructure. Due to security purpose of the organization these users account registries are not shareable [36]. It is not acceptable to replace existing reliable and secure authorization infrastructure which is managed, supported and deployed with single model security mechanism [36]. In order to have successful grid security architecture there is need of mechanism to which perform integration between existing technologies, hosting environment and across model platform [36].

Interoperability Challenge:

In the grid environment the services are extend over the hosting environment and domains. In order to work properly they need to interact with each other [36]. Due to this the need of interoperability increases at different level such as [36].

- **Protocol level:** In this level such mechanism is needed that allows messages are exchanged in domain. For example this can be achieved with SOAP/HTTP [36].

- **Policy level:** In this level such mechanism is adopt in which each party explain polices to other parties which it may use in secure conversation [36].

- **Identity level:** In identity level such a mechanism is needed which helps to identifying users of one domain in other domain [36].

Trust relationship challenge:

Due to the dynamic nature of the grid the request of service may be span on more then one security domain. To get the out come of the end-to-end traversal this domain trust relation can play important role [36]. On every request trust establishment may be evaluated dynamically or one time activity per sessions [36].

Following is the categorized solution of the security challenges.

- The solution for integration is to have an extensible architecture in which interfaces should be abstract and also utilized the existing services where it needed [36].

- To have interoperability solution there must be some mechanism to invoke the services that are hosted in different virtual organization and security mechanism [36].

- In grid environment to enforce and manage the trust polices there must be a mechanism for having the solution for this security challenge [36].

In figure 4.9 depicts the dependencies of these categories

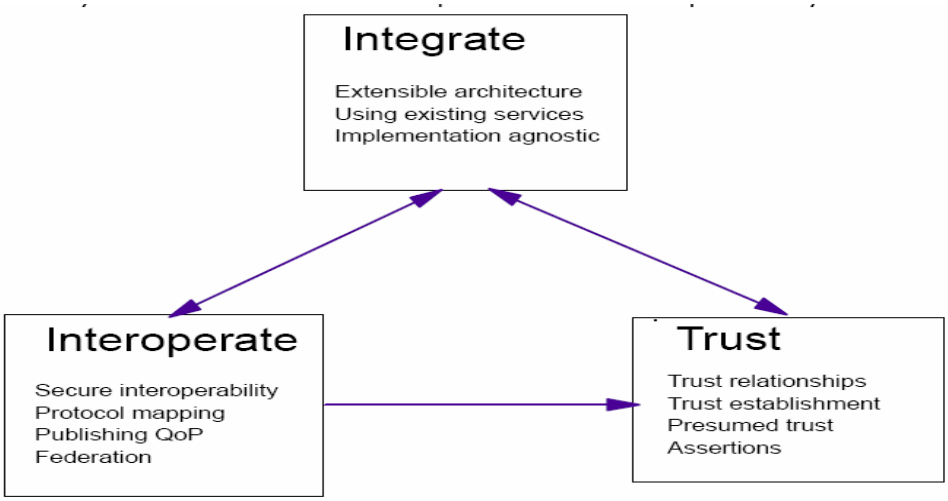

Figure 4.9 : Dependencies of security challenges solution [36]

The security model must address the following security requirement [36].

- **Authentication:** The authentication mechanism of the security model must provide ways to convey the specific mechanism which is used in any given authentication operation and also for multiple authentication mechanism provide

plug point [36]. The mechanism used in authentication may be custom based or an industry based technology [36].

- **Delegation:** The security model must allow the delegation in polices to be specified and access right from requestors to services [36].

- **Single Logon:** In security model when the authentication performed successfully then there is need of mechanism which must give relieve to entity for further authentication for accessing the resources in grid [36]. When subsequent is access made by entity then there is need further re-authentication for security [36].

- **Renewal and credential lifespan:** In some cases of grid environment the jobs take more time then initially user delegated credential. In these scenarios the security model must have the ability to refresh the job credential and also have the ability to prior notify about expiration of the job credential to user [36].

- **Authorization:** In security model the OGSA services controlling access are allowed on the basses of authorization policies. This mean under what condition the entity can access the service [36].

- **Manageability:** In security model must have the ability to explicitly determine the need of manageability in security functionality [36].

- **Firewall traversal:** Today the firewall becoming barrier for dynamic, cross domain grid computing. It might possible in near future the firewall disappears. In this scenario security model must have the capabilities to overcome the need of the firewall [36].

- **Privacy:** The security model must have the ability to enforce privacy polices by allowing the service provider and requestor to define these policies [36].

- **Confidentiality:** The security model must have the ability to protect confidentiality during in the flow of the message or documents in OGSA compliant infrastructure and in underlying communication mechanism [36].

- **Integrity of message:** There is an ability to detect changes made by unauthorized way in message and documents in security model. Integrity can be achieved by determining the quality service polices [36].

- **Assurance:** The security model provides the way to increase the security assurance level which is expected by the hosting environment [36].

- **Policy exchange:** To establish the negotiated security context between them the service provider and requester are allowed to exchange the dynamically security polices among other [36].

- **Secure logging:** The security model must have this ability to address the requirement for auditing, notarization and non-repudiation [36].

The following is security model for the proposed solution for the Protection component of the RIS.

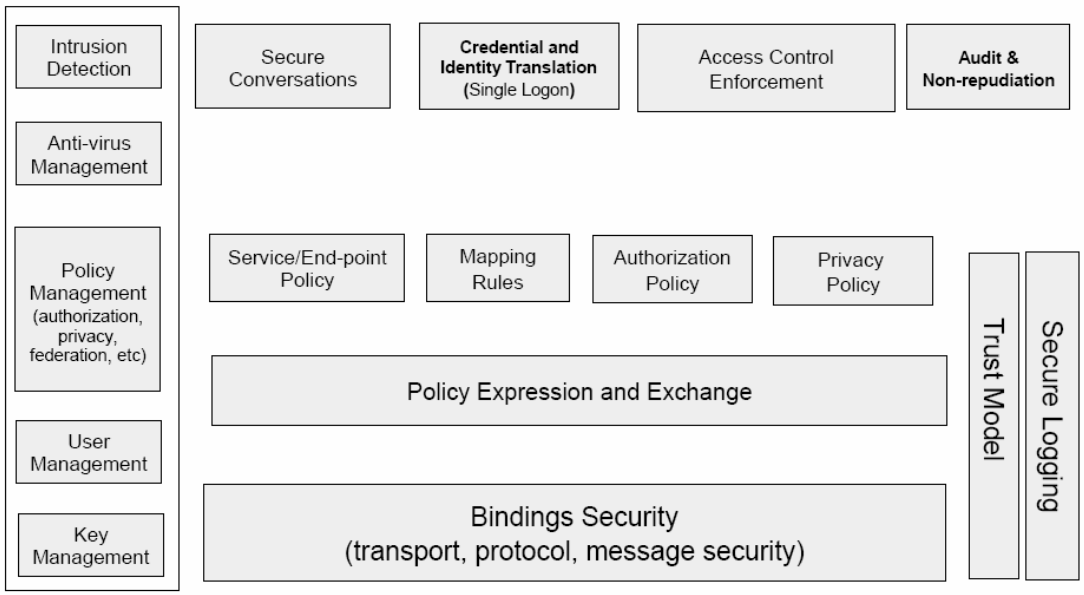

Figure 4.10: Security model [36].

- **Binding security:** The message format and security characteristics of the associated protocol are things on which security binding is based on. The minimum suitable authentication, confidentiality and integrity can be achieved by taking care of the security requirement in binding process, when new message format and protocol are introduced [36].

- **Policy and expression exchange:** To exchanging the policy between participating end points it requires the policy expression and exchange facility that will address requirement for grid security. It is also play a critical part to achieve secure association between the points and securing infrastructure of the Open Grid Service Infrastructure (OGSI) [36].

- **Secure association:** To establish the identity of a requestor to the service provider the secure association is required .This might able the service provider to satisfy the requirement to authenticate the identity on other end and then privacy and authorization policies based on the established identity is enforced [36].

- **Identity and credential mapping /translation:** The function which is directed by the corresponding policies is implemented by the mapping/translation components at this layer [36].

- **Authorization enforcement:** In authorized policies those policies are also to be included which is required by grid security model. For security model the authorization is key part. To make its own access decision each domain will typically have its own authorization service [36].

- **Privacy enforcement:** To enforce the privacy policies in grid environment the Web services (WS) policy, WS-privacy and WS-Authorization should adopt by security model [36].

- **Trust:** WS-trust specification on which the grid trust model is based. Trust policies which are define either a –prior or dynamically can establish and enforce the trust [36].

- **Secure logging:** The secure logging functionality explicitly called grid security model which are necessary for foundation of many higher level audit related services [36].

- **Management of security:** The management of security model includes the key management of cryptographic function, privacy and trust policy management, user registry management, management of mapping rules which helps to enables federation and authorization [36].

4.4 Detection component of RIS

The need of detection component arises when the protection component is compromised by malicious intent or when the RIS performance is decreased. Responsibilities of detection component of RIS is to find the causes that decreasing of performance of the RIS. The detection processes are initiated by the monitoring component of the RIS. For doing this RIS monitoring model use monitoring data which are gathered by performing monitoring process in RIS.

As the RIS is emerging; technologies and infrastructure on RIS based is grid. The grid system does not have such mechanism that fulfills the requirement of RIS detection component. Following are some basic requirements that detection component of RIS must have.

- To perform detection processes the detection component used such a mechanism in which detection algorithms and heuristics are used to find the causes which are decreasing the performance of the RIS.

- The detection mechanism must have the capability to analyze the monitoring data. These monitoring data can be received from RIS knowledge base component.

- The detection mechanism must be intelligent enough to perform operation after the analysis of the monitoring data .For example if it is found from monitoring data that some operations of RIS is not performing well then detection mechanism must have the capability to detecting the defects of the these operation.

- The result of detection process must be stored in knowledge base component of RIS which helps the protection system to improve its protection mechanism and restoration system in restoration process.

- The detection mechanism must have the ability to initiate the restoration processes to perform restoration after finding the defects.

Making the detection mechanism powerful helps us to maintain the performance of the RIS because as fast as detection process detects the fault; it helps the RIS to come out from the catastrophic situation fast by starting the restoration process as early as possible.

4.5 Restoration component of RIS

After the protection and detection the restoration component RIS plays important role to bring out RIS from catastrophic situation. The restoration mechanism must have the

Following basic requirement in order to perform the restoration process in RIS.

- Increase the optimization of the RIS
- Perform self configuration in RIS
- Perform self healing in the RIS

The explanation of the above requirement of restoration is following

Self healing: When fault occurs then restoration component must have the ability to ensure the effective recovery from fault. This can be done by the help of detection component of RIS to properly determine the fault and provide this information to restoration component to perform restoration with the minimal disruption of the RIS operation [37].

Optimization: The restoration component must have the ability to perform optimization of RIS. This can be done by knowing the ideal performance of RIS and comparing the current performance of RIS with ideal performance in specified time interval [37]. Monitoring data of the monitoring component of RIS helps to performing this task.

Self Configuration: The restoration component must have the ability of self configuration by adopting the changes according to the current scenario [37].

The restoration component processes are initiated by detection component of the RIS. The restoration component gets the data of the faulty area from the knowledge base and when it performs the restoration; it also updates the knowledge base by giving the data about the faults against which it restored. This data helps in optimization process by determining the performance of the RIS after the restoration process.

4.6 Knowledgebase component of RIS

The knowledgebase component is a most essential part of the RIS. Without this component the other components of the RIS do not able to perform their operation successfully .The reason is that all the data require by these components are store in knowledgebase component of the RIS. Think knowledge base as a big database which contains all the data which are generated by the components of RIS and this data is also utilized by these components in order to perform their functions successfully. Following are some essential types of data which is stored by knowledge base.

- **Monitoring data:** Monitoring data is one of the data types which are generated by monitoring component during monitoring process. This type of data is utilized by almost every component of the database.
- **Vulnerabilities information:** The vulnerabilities information is one of the most important information which is stored in the knowledgebase of the RIS. This information is needed by the other component of the RIS such as protection, restoration and detection to perform their operation. The protection component needs these vulnerabilities information in order to protect the RIS from the

attacks which are based on these vulnerabilities. The vulnerabilities information is needed by the detection component in order to detect the malicious intents which are based on these vulnerabilities. The restoration component uses these vulnerabilities information in order to perform their restoration operation bases on these vulnerabilities solution. But there is need of standardizing the vulnerabilities in order to store information of about these vulnerabilities. The MITRE cooperation of US gives the concept of standardizing the information vulnerabilities by the name of CVE (Common Vulnerabilities and Exposures) [38]. CVE is funded by National Cyber Security Division of the US Department of Homeland Security [38]. The aim of CVE is to developed the vulnerability dictionary in order to standardized the public know vulnerabilities name and exposures. There is also vulnerabilities database organized NIST(National Institute of Standard and Technologies) by the name of NVD (National Vulnerabilities Database) which is also funded by National Cyber Security Division of the US Department of Homeland Security [39]. The CVE naming standard one which NVD base and synchronized [39].The vulnerabilities information can be gather by the cooperation these organization.

- **Restoration, detection and protection techniques:** The knowledgebase also have this type of data which helps by giving the information to restoration, detection and protection component of the RIS in order to perform their processes. This data might contain instruction and method for techniques which are implied by these components.

- **Some extra data:** This type of data is also stored in knowledgebase which is produced by the components of the RIS in order pass the data to other component to perform their operation and perform communication between them.

5 DISCUSSION

In this chapter the research work is summarized and discussed.

5.1 Gap identification

The role of the information systems has become vital in today's world and their application area has crossed all the geographical boundaries. Information systems are providing services in public & private sector by performing business & management operations and help in scientific and academia research. These information systems are important part of critical infrastructures and they can not be left fragile in catastrophic situations. The gap in this area can be filled by having a system which has the ability to withstand and keep performing its operations under any type of stress/attack with the resources available; such an ability of information systems is resilience. The motive behind conducting this research is to define resilient information system and design it's meta-model.

5.2 Findings

To address the research question of this thesis the first thing was to define resilient information system which is defined and explained in section 3.1 and 3.2. According to definition the resilient information system (RIS) must have to following characteristics.

- Self protection
- Self healing
- Self configuration
- Self optimization

Self protection is a RIS characteristic used to protect against attacks and in future give protection against new found attacks which were not protected in past. Self healing is a RIS characteristic used to identifying and fix frequently accruing problems. Self configuration is a RIS attribute using which information systems can modify its present configuration in reaction to the current workload. Self optimization is a RIS characteristic used to analyze the situation, determining the goals and adjustment of the behavior. The meta-model of RIS is defined in section 3.3 to achieve these characteristics. The components of meta-model are as follows:

- Monitoring
- Protection
- Detection
- Restoration
- Knowledgebase

To obtain the above mentioned characteristics of RIS the components of meta-model of RIS are used in the following manner.

Self Protection: Protection component of RIS works at the entry points. The protection component gets the information from knowledgebase component of the RIS. All the messages coming to RIS will be monitored. Message authentication is performed in monitoring by verifying its source. The message is then forwarded to the protection component to verify the signature of the message. If the signature is valid then it will be allowed to perform its operations. If some malicious message compromise the protection component then detection component will be initialized by monitoring component to detect the malicious message. The monitoring component will identify the problem by performance monitoring. The detected malicious message will update the knowledgebase for up-gradation of the protection component.

Self Healing: This characteristic helps to heal the RIS from the catastrophic situation to bring back to its normal working state. The process of restoration component is initiated by the detection component at detection of the faulty part and this information is provided to restoration component to perform restoration on these parts. The information to perform restoration is gathered from the knowledgebase component of the RIS.

Self Configuration: The monitoring component will check whether the workload has increased at some component with the help of detection component. The detection component will check for the parts where workload has increased. The restoration component will re-configure the system for load balancing on the overloaded points.

Self Optimization: Optimization will be performed with the help of restoration, detection and monitoring component. Monitoring component initiates the detection component when it found the RIS is under performing. Detection component will find the reasons for this under performance. This information will be passed to restoration component to fine tune or optimize the performance of RIS.

Grid as RIS Infrastructure: RIS requires cheap and extra resources for performing the operations. RIS also require interoperability, reliability, flexibility, manageable infrastructure and open & standard architecture. This can easily be achieved by implementing VO structure of the grid. The grid infrastructure is most appropriate solution for RIS as the grid infrastructure has more resources and is operated on Internet and its reliability can easily be achieved. The OGSA based grid implementation fulfills the requirement of open and standard architecture. As grid is selected as infrastructure of RIS, the components of RIS are integrated and implemented in grid.

6 FUTURE WORK

As the research area is new and there is a lot of work to do and test in the research discipline. We have tried to categorize the research areas for future studies.

6.1 Research on RIS Meta-model testing and implementation

- Due to the limitations of time and resources the meta-model of RIS is not implemented. As the proposed RIS meta-model is designed for implementation on grid infrastructure. A research can be conducted to improve the grid technology to support RIS meta-model components.

- A need is to conduct a research for the development of a framework for the implementation of RIS on grid infrastructure.

- A research can be conducted to find more components for improvement in the meta-model to strengthen resiliency of information systems.

- A research can be conducted for RIS tools development.

6.2 Research on RIS components

- A research can be conducted to improve the components (monitoring, protection, detection, restoration and knowledgebase) of RIS.

- A research can be conducted to develop monitoring mechanism according to RIS characteristics.

- A research can be conducted to develop monitoring tools.

- A research can be conducted to improve the protection mechanism according to RIS characteristics.

- A research can be conducted to develop protection tools.

- A research can be conducted to improve the detection mechanism according to RIS characteristics.

- A research can be conducted to develop detection tools.

- A research can be conducted to improve the restoration mechanism according to RIS characteristics.

- A research can be conducted to develop restoration tools.

6.3 Other research areas

- A research can be conducted to make and standardize a scale to measure resiliency of a RIS.

- A future research can be conducted for the implementation of the proposed meta-model to add resiliency components in information systems at development phase.

- A research can be conducted to increase the intelligence of RIS.

REFERENCES

[1] Cabinet office, Central Sponsor for International Assurance (2007) "*Protecting our information systems*". http://www.cabinetoffice.gov.uk/csia/documents/pdf/CSIA_booklet.pdf. Last Accessed: 2007-3-14

[2] Nemeth C, Nunnally M, Connor M, Cook R (2006) "Creating Resilient IT: How the Sign-Out Sheet Shows Clinicians Make Healthcare Work". Department of Anaesthesia and Critical Care, University Of Chicago, Chicago, IL, USA. http://www.ctlab.org/documents/AMIA-86319-Nov2006.pdf.Last Accessed: 2007-03-14

[3] Spafford G (2005) "*Creating a Resilient IT System*". http://itmanagement.earthweb.com/netsys/article.php/3509266.Last Accessed: 2007-3-14.

[4]Hornby A S (2005) *"Oxford advanced learner's dictionary"*. Oxford University press. pp 1291 and 1557

[5]Varga M (2003) *"Zachman Framework in Teaching Information Systems"*. In Proceeding of 25th International Conference on Information Technology Interfaces, IEEE, PP 161-166.

[6] Walker B, Holling C. S, Carpenter S.R, Kinzig A (2004) "*Resilience, adaptability and transformability in social–ecological systems*". Ecology and Society. http://www.ecologyandsociety.org/vol9/iss2/art5/. Last Accessed: 2007-3-18.

[7] Macquarie R, Palo M, Surname A (2002) "*Network Resilience: Resilient computing through network design and workforce mobility*", IBM global services. http://www-935.ibm.com/services/us/gn/pdf/network_resiliency_white_paper_gw510-3036-01f.pdf. Lat Accessed: 2007-3-20

[8] Kephart J.O, Chess D.M (2003) *"The vision of Autonomic Computing"*. Computer, IEEE Computer Society, Volume 36, Issue 1,PP 41-50.

[9] Lea A (2004) *"Psychological models in autonomic computing systems"* .In Proceedings of 15th International Workshop on Database and Expert Systems Applications, IEEE, PP 747-751

[10] Mowbray M, Williamson M (2003) "*Resilience for Autonomous Agents*". Internet System and Storage Laboratory, Hewlett-Packard Laboratories Bristol. http://www.hpl.hp.com/techreports/2003/HPL-2003-210.pdf. Last Accessed: 2007-04-21

[11] Dalziell E.P, MacManus S.T, "*Resilience, Vulnerabilities, and Adaptive Capacity: Implication for System Performance*". Department of Civil Engineering, University of Canterbury, New Zealand. http://www.ifed.ethz.ch/events/Forum04/Erica_paper.pdf. Last Accessed: 2007-04-21

[12] DTI Technical Report (2006) *"Information Security Breaches Survey 2006"*. http://www.dti.gov.uk/files/file28343.pdf , Last Accessed: 2007-04-22

[13] A brief history of the grid,
http://gridcafe.web.cern.ch/gridcafe/Gridhistory/ancestors.html , Last Accessed:
2007-04-30
[14] Foster I, Kesselman C, Tuecke S (2001) *"The Anatomy of the Grid: Enabling Scalable Virtual Organizations"*. International Journal of High Performance Computing Applications, Vol. 15, No. 3, 200-222

[15] What is grid computing,
http://www-03.ibm.com/grid/about_grid/what_is.shtml . Last Accessed: 2007-05-01

[16] Ian Foster (2002) *"What is the grid? A Three point checklist"*. Argonne National Laboratory & University Chicago.
http://www-fp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf .Last Accessed:
2007-05-01

[17] Grid architecture,http://gridcafe.web.cern.ch/gridcafe/gridatwork/architecture.html .Last visited:
2007-05-01

[18] Ferreira L ,Berstis V , Armstrong J , Kendzierski M, Andreas A, Takagi M , Bing-Wo R, Amir A , Murakawa R, Hernandez O , Magowan J, Bieberstein N (2002) *"Introduction to Grid Computing with Globus"*. Red books, International Technical Support Organization, IBM, First edition .

[19] The Globus Alliance,
http://www.globus.org/ . Last visited: 2007-05-03

[20] Foster I, Kesselman C, Nick J.M, Tuecke S (2002) *"Grid Services for Distributed System Integration"*. Computer, IEEE Computer Society, Volume 35, Issue 6, PP 37-46

[21] Foster I, Kesselman C, Nick J.M, Tuecke S (2002) *"The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration"*. Globus Project.
http://www.globus.org/alliance/publications/papers/ogsa.pdf .Last Accessed: 2007-05-04

[22] Kishimoto H, Treadwell J (2005) *"Defining the Grid: A Roadmap for OGSA™ Standards*". Open Grid Services Architecture Working Group.
http://www.ogf.org/documents/GFD.53.pdf . Last Accessed: 2007-05-04

[23] Kishimoto H (2006) "Defining the grid: Open Grid Services Architecture". Global grid forum.
http://www.ggf.org/GGF17/materials/316/OGSA%20keynote%2020060510.ppt.
Last visited: 2007-05-04

[24] Imamagic E, Radic B, Dobrenic D (2005) *"CRO-GRID Grid Monitoring Architecture*". In Proceeding of 27th International Conference on Information Technology Interfaces, PP 65-72.

[25] Cooke A, Gray A.J.G, Ma L, Nutt W, Magowan J, Oevers M, Taylor P, Byrom R, Field L, Hicks S, Leake J, Soni M, Wilson A, Cordenonsi R, Cornwall L, Djaoui A, Fisher S, Podhorszki N,Coghlan B, Kenny S ,Callaghan D O (2003) *"R-GMA: An Information Integration System for Grid Monitoring"*. Lecture Notes in Computer Science, Springer-Verlag Berlin or Heidelberg, Volume 2888/2003.

[26] Tierney B, Crowley B, Gunter D, Lee J, Thompson M (2001) *"A Monitoring Sensor Management System for Grid Environments"*. In Journal Cluster Computing, Springer Netherlands, Volume 4, Number 1 / March, 2001 .
http://www.springerlink.com/content/g30700611l6hu143/ . Last visited: 2007-05-05

[27] Cooke A, Nutt W, Magowan w, Taylor P, Leake J, Byrom R, Field L, Hicks S, Soni M ,Wilson A, Cordenonsi R, Cornwall L, Djaoui A, Fisher S, Coghlan B,Kenny S,Callaghan D.O ,Ryan J (2003) "*Relational Grid Monitoring Architecture*". UK e-Science All Hands Conference, Nottingham.
http://www.r-gma.org/pub/ah03_148.pdf . Last Accessed: 2007-05-08

[28] R-GMA in 5 minutes.
http://www.r-gma.org/fivemins.html . Last Accessed: 2007-05-08

[29] Monitoring Agents using a Large Integrated Services Architecture.
http://monalisa.cacr.caltech.edu/monalisa.htm . Last Accessed: 2007-05-08

[30] Legrand I.C,Neman H.B,Voicu R,Cirstoiu C, Grigoras C, Toarta M, Dobre C (2004) *"MONALISA: An Agent Based, Dynamic Service System To Monitor, Control And Optimize Grid Based Applications"*. CHEP 2004, Interlaken, Switzerland.
http://monalisa.cacr.caltech.edu/documentation/monalisa_chep04.pdf . Last Accessed: 2007-05-08

[31] Legrand I.C (2005) *"MonALISA - An Agent Based, Dynamic Service System to Monitor, Control and Optimize Distributed Systems"*. ICFA WORKSHOP, Daegu.
http://monalisa.cacr.caltech.edu/documentation/MonALISA_Daegu_May05.pdf.
Last Accessed: 2007-05-08

[32] Massie M .L, Chun B. N, Culler D. E (2004) *"The ganglia distributed monitoring system: design, implementation, and experience"*. Parallel Computing; 30(7), 817–840.
http://ganglia.info/papers/science.pdf . Last Accessed: 2007-05-08

[33] Wolski R, Spring N T, Hayes J (1999) *"The network weather service: a distributed resource performance forecasting service for metacomputing"*. In Journal of Future generation Computer Systems, volume 15, version 5, PP 757-768(12).

[34] Padgett J, Djemame K, Dew Peter (2005) "*Grid Service Level Agreements Combining Resource Reservation and Predictive Run-time Adaptation"*. In proceeding of School of Computing, University of Leeds, LS2 9JT, United Kingdom.

[35] Padgett J, Djemame k, Dew Peter (2005) *"Grid-Based SLA Management"*. Lecture Notes in Computer Science Springer Berlin / Heidelberg, Volume 3470/2005.

[36] Nagaratnam N, Janson P, Dayka J,Nadalin A, Siebenlist F, Welch V ,Foster I, Tuecke S (2002) *" The Security Architecture for Open Grid Services"*. Version 1.
http://www.di.unipi.it/~coppola/GRIDsem/OGSA-SecArch-v1-07192002.pdf .
Last Accessed: 2007-05-13

[37] Sterritt R, Bustard D (2003) *"Towards an Autonomic Computing Environment"*. In Proceedings of the 14th International Workshop on Database and Expert Systems Applications, PP 694-698.

[38] Common vulnerabilities and exposures.
http://cve.mitre.org/ . Last Accessed: 2007-05-15

[39] National Vulnerabilities Database.
http://nvd.nist.gov/nvd.cfm . Last Accessed: 2007-05-15

[40] Encylopedia Britinica *"Information System"*.
http://www.britannica.com/eb/article-9126502/information-system .
Last Accessed: 2007-05-14.

[41] Mark Kelly (2005), *"Components of Information Systems"*. Lecture notes in computer science Mckinnon secondary college.
http://www.mckinnonsc.vic.edu.au/la/it/ipmnotes/systems/components.htm#equip .
Last Accessed: 2007-05-14.

[42] Information and Communication Technologies.
http://cordis.europa.eu/fp7/ict/ . Last Accessed: 2007-05-08

[43] EURESEARCH.
http://www.euresearch.ch/fileadmin/documents/PdfDocuments/Callfiches/Call1_fiche_ICT.pdf. Last Accessed: 007-0508

[44] IBM *"Autonomic computing"*.
http://www.research.ibm.com/autonomic/overview/ .Last Accessed: 2007-05-10

[45] Kentucky Center for Resilient Information Systems.
http://protocols.netlab.uky.edu/treasury/. Last Accessed: 2007-05-18

[46] Creswell J (2002) *"Research Design: Qualitative, Quantitative and Mixed Method Approaches"*. Sage Publications Ltd.

[47] NetMBA, Maslow's hierarchy of needs
http://www.netmba.com/mgmt/ob/motivation/maslow/. Last Accessed: 2007-05-20

[48] Predictive Self-Healing in Solaris 10 Operating System, http://Www.Sun.Com/Bigadmin/Content/Selfheal/Selfheal_Overview.Pdf. Last Accessed: 2007-05-20

[49] Joseph R.K (2003) *"Global Information Systems Threats"*. Seminar in global strategic information systems, Infs 6750.
http://www.savageideas.com/downloads/mba/Global_Information_Systems_Threats.pdf.
Last Accessed: 2007-05-20