Chapter II

Review of Literature
CHAPTER 2

REVIEW OF LITERATURE

2.1 Introduction

The last decade has seen a considerable increase in commodity computer and network performance mainly, as a result of faster hardware and more sophisticated software. Nevertheless, there are still problems, in the fields of science, engineering and business, which cannot be dealt effectively with the current generation of supercomputers. In fact, due to their size and complexity, these problems are often numerical or data intensive and require a variety of heterogeneous resources that are not available from a single machine. A number of teams have conducted experimental studies on the cooperative use of geographically distributed resources conceived as a single powerful computer. In this chapter, a related work is presented that represents the problems and issues related to advance reservation of the resources and scheduling for a specific class of grid platforms.

2.2 Applications of ICT in government sector

With increasing awareness about ICT [95], information systems continue to proliferate all over the world at a breathtaking pace. ICT has made a profound impact on our lives and times. The emergence of web-based delivery system has increased the relevance of these tools in the collection, collation and sharing of
information at low cost. Above all, this technology holds out the promise of transforming societies in the developing world. For instance, it can help social sectors like education, health, rural development and other areas to reach out and deliver the goods and services to distant consumers. Hence, it is important to maximize the use of ICT tools and applications for successful transformation of government delivery systems with minimized cost and maximized speed.

New information and communication technologies can also make significant contribution towards achievements of good governance goals. But, government programmes have a number of characteristics that make their information needs different from those of the other corporate and business organizations.

These include:

- A demand-driven approach, often with limited management capacity.
- A large number of programme implementing agencies, often with limited management capacity.
- A very large number of beneficiaries, with disbursements and delivery taking place daily.
- A variety of procurement methods, with preponderance of community-based procurement.
- Operations in different departments require wide range of skills, motivation approaches, service providers and contractors.
- Frequent decentralization and centralization of management activities.
• Frequent contracting of important infrastructure activities.
• An anti-poverty agenda requiring the selection and monitoring of specific indicators for evaluating impact.
• Multiple stakeholders (citizens, communities, government, political executives, bureaucracy, employees, contractors, non-governmental organizations etc.), with distinct information needs.
• Political visibility making transparency and efficiency is more important. Therefore, the strategies and implementation task of using ICT in government sector demands advanced vision, skills at higher authority level and cooperation from all.

2.3 Issues of e-governance

Governments all over the world are trying to utilize IT for various purposes [2] [14]. The initial motivation usually comes from the need to improve efficiency of processes in the government. This may be concurred or followed by the second step comprising re-engineering of the processes.

An e-government application brings government into more than efficient and effective government in delivery of services and internal administration and communication. E-government brings countries and concomitantly its citizens, into the international stream.

This brings cultural change within the government and in the overall society. Increased links by government and the citizen alike, to international web sites,
news groups, chat rooms, email communications, building of international networks through the use of email and the powerful wireless technologies all contribute to an emerging linked world that operates continuously.

The great potential of new technologies for enhancing democracy without supposing that either this is yet happening or that it will necessarily happen without a lot of sustained effort that should be considered.

There are very considerable opportunities but there are too risks. Those can most effectively be minimized if we first open our eyes to both the benefits and the potential costs of the change in the culture of democracy that the new technologies are bringing about and are likely to continue to bring about.

Gartner [77], an international consultancy firm, has formulated a four-phase e-governance model. This can serve as a reference for governments to position where a project fits in the overall evolution of an e-governance strategy.

According to Gartner, e-governance will mature to the following four phases viz., Information, Interaction, Transaction and Transformation.

E-government is itself a process or a means to an end, rather than an end in itself. E-government is still in the earliest stages of development and promises to evolve with advances in technology and increased acceptance and trust in electronic communications. Initial forays into e-government initiatives have focused mostly on providing enhanced access to information and basic services.
Although the full transformative effects of e-government remain largely unrealized at this time, the rapid growth in interest and resources dedicated to e-government initiatives may contribute to swifter changes.

A set of motivation may come from the need to provide various social services to citizens for improving their quality of life. A parallel set motivation may be used to strengthen the democratic foundations of governance (opinion polls, voting etc.).

These social services and democratic enablement corresponds to the new activities that become economically viable due to the altered cost structure and use of information technology.

Government grows up exponentially once e-governance is introduced. This is because the ease of transactions introduced by e-governance would encourage citizens and businesses to have more transaction with the government. In order to manage this enormous amount of data such that, system performance does not get degraded and such that the system is scalable and there is a need to automate the record management functions.

Although there exists solutions for document archival system based on network-centric groupware, they do not address the issue of archival based on policies of multiple. While initiatives have been emanating from various directions, they are often at cross-purposes and so repetitive and wasteful.

Many governments are struggling with massive internal re-engineering in relation to technology to improve efficiencies and deliver government services
electronically to citizens. Thus e-governance would require setting up of LANs/Intranets/Extranets, along with data warehousing, administrative process revamp, training and capacity enhancement in government organizations.

Economic benefits accruing to governments from technology integration have historically been viewed in terms of cost savings and return on investment from specific projects. Effective use of technology is therefore a key to successful deployment of electronic governance.

Accordingly, critical issues need to be clearly delineated and solutions worked out. Whatever may be the middleware chosen, the system should be highly reliable, must have capability for future expansion, management with ease, and high performance. But above all should have very good security in transactions. In order to achieve this system like e-governance, must have some built-in services will not only make the applications to have smooth interactions but also make the users feel that the entire operation is user friendly. So that e-government is about transformation that helps citizens and businesses find new opportunities in the world’s knowledge economy. It holds great potential.

Yet, if e-government is not part of a larger program for reform—reforming how government works, manages information, manages internal functions, serves citizens and businesses—then it may not produce all the benefits expected from the time and money invested. Use e-government to rethink the role of government. Use it as a tool to further economic development and good governance.
The infrastructure needs and issues of e-governance cannot be generalized. The e-governance infrastructure varies according to the governance objectives and needs of the society. The architecture options are decentralized architecture, hybrid architecture and centralized architecture.

Social Issues

Social issues play an important role in the success and failure of the e-governance efforts. The social changes will happen on a constant basis. E-governance is not just a technology-oriented effort rather it aims at creating the concept of social change.

This change can be at the level of the administration, politics, social, economical or individuals, the society is a dynamic body. Civil society as a concept can be difficult to understand but although there are number of conflicting definitions there is widespread agreement during the past few decades, civil society receded and political/commercial society advanced in terms of their impact on people’s lifestyles.

Civil society is based on voluntary participation, whereas the state is based on coercion. It can be argued that political society in pseudo-democracies has a leg in each camp and consists of competing self-serving organized gangs intent on gaining control of state apparatus. Government is institutionalized coercion and must be so, since we have governments in order, essentially, to protect civil society from internal and external predators.
Some define civil society to include only non-profit organizations, others define it to include only self-organizing communities of common interest, others apply the descriptor to all forms of non-governmental cooperation including big business, while yet others define it to exclude all forms of institutionalized human activity.

Sorting the range of different definitions in use isn't helped by the fact that what once distinguished government from the market was coercive force, present in the one and absent in the other.

However, capital has steadily become more concentrated in fewer hands, and governments have relied more and more on big business to finance election campaigns. Therefore quite apart from governments' innate propensity to drift towards crony-capitalism, the above differentiating line has faded almost to non-existence.

The basic objective of the e-governance exercise is to create a transparency among the various stakeholders of the society namely the government, administration, business and common masses. But most of the societies are fearful of the transparency due to the possible pilferages that might occur due to this factor of transparency.

Various stakeholders have their own personal motives, which can act as barrier in successful implementation of an e-governance structure in place. This gets more critical in the under developed nations where typically the administration and the politics tend to thrive on the red tapism, bottlenecks and
corruption. With an e-governance structure in place they become vulnerable to the checks and balances. Often it is observed that the powerful stratum of the society tend to adopt a non-cooperative approach towards the entire issue. The social awareness is thus an extremely important component to determine the success of e-governance endeavor.

Here it states that e-governance is not just a technical issue rather it’s a social issue. The worthiness of the initiative lies in its usage, and as a first point to be noted is the need to use local languages in the IT implementation process.

However, in a broader context it is essential that a clear strategy to have access to local level databases maintained in regional languages as well as to use appropriate interfaces to aggregate such data. Here, a focused effort would have to be made in improving the technologies for transliteration.

Similarly, an effort would have to be undertaken for perfecting the Optical Character Recognition (OCR) technology for local languages. This is critical, as an effective OCR technology is required to convert the data that is scanned and stored in local languages into meaningful and workable databases.

This could also help in substantially simplifying the efforts at content generation and data warehousing. Government data is huge and complex. There exists a large inventory and repository of appropriate technological platforms with which to perform public functions in a cost and time effective manner.

However, common platform availability is an issue in this. A major portion of the government data is in the form of what we categorize as legacy data.
Integration of the legacy data into the e-governance initiative is not only time consuming but also expensive. Software is not the only issue that poses a challenge to the e-governance initiatives.

Only Computerization of government work will lead to automation of existing manual work. There is a real danger of investments being made in hardware, application software and even in training. But still the issue of prompt and regular service, updating of data, not being effectively addressed.

The e-governance initiative has to address the technological problems objectively by identifying functional areas in every government organization to initiate the process of computerization and choosing appropriate hardware platform and software applications for cost effective delivery of public services.

Simply drafting a law or issuing an order from political leaders, as with all reforms, cannot achieve e-government. It requires changing, how officials think and act, how they view their jobs and how they share information between departments.

Different level and model of computer applications for ministry, department, directorate and offices in urban and rural areas that could be suitably customized as per location and work specific requirements should be created.

The issue of management of change, which has to be made quite rapid at times, is the other major issue to be addressed while implementing IT for good governance. Changes in the decision making procedures, processes, delegation of
power etc. would lead to mandatory organizational and institutional changes affecting both peoples and methods at all the bridges of delivery chain.

For this, acceptance of changed processes would have to be properly understood, accepted, adopted and improved to enable full advantage of the Information technology being adopted for e-governance.

The e-governance initiatives are likely to induce a great degree of change in how governments are functioning today. Change management is a subject matter that is currently sweeping public and private sector organizations. It is not new but has been evolving as a discipline over the past decade.

Change management is addressing the changes being faced by modern public administrators both internally and externally. In the past three decades governments have had to contend and deal with many societal evolutions that have significantly altered the way public administrators function in their jobs.

Globalization, free trade, the increased movements of people and goods, internationally, on a scale never before known, new attitudes of a citizenry that is expecting more from governments, changing social attitudes and new technologies, especially information and communication technologies, have all contributed to the challenges now facing politicians, executives and managers alike.

Change management, in the context of e-government is all about how members of the public service make the transition from the traditional approaches to management, Pre-Information and Communication Technologies era, to new
means of administering in new and evolving environments. The latter principle is one of the many ways in which change management is different for the public sector.

Although e-government encompasses a wide range of activities and actors, three distinct sectors can be identified [3] [4] [6]. These include government-to-government (G2G). Government-to-business (G2B) and Government-to-citizen (G2C). Some observers also identify a fourth sector, government-to-employee (G2E).

**Government-to-Government (G2G):** In many respects, the G2G sector represents the backbone of e-government. Some observers suggest that governments at all levels must enhance and update their own internal systems and procedures before electronic transactions with citizens and businesses can be successful. G2G e-government involves sharing data and conducting electronic exchanges between governmental actors. This involves both intra and inter-agency exchanges at the national level, as well as exchanges between the national, provincial, and local levels. G2G applications constitute the maximum chunk of the e-governance functions.

**Government-to-Business (G2B):** Government-to-Business (G2B) initiatives receive a significant amount of attention, in part because of the high enthusiasm of the business sector and the potential for reducing costs through improved procurement practices and increased competition. The G2B sector includes both
the sale of surplus government goods to the public, as well as the procurement of goods and service. Although not all are directly dependent on the use of information technology, several different procurement methods are used in relation to the G2B sector. Tendering, auctions and information and administrative management tools form the core of this sector.

Government-to-Citizen (G2C): The third e-government sector is Government-to-Citizen (G2C). G2C initiatives are designed to facilitate citizen interaction with government, which is what some observers perceive to be the primary goal of e-government. These initiatives attempt to make transactions such as renewing licenses and certifications, paying taxes, and applying for benefits, less time consuming and easier to carry out. G2C initiatives also often strive to enhance access to public information through the use of dissemination tools, such as Web sites and/or kiosks. Another feature of many G2C initiatives is the effort to attenuate the agency-centric and at times, process-laden nature of some government functions.

The basic aim of any e-governance initiative is aimed at good governance and government citizen interface is thus the most critical aspect of the entire exercise. Digitization of all public domain information like official gazette notifications, acts, rules, regulations, circulars, policies and program documents are some of the important portions of this exercise.
To implement an e-governance model, this should be composed of six key elements:

- **Service provision** – A set of high-quality services delivered with innovative methods to user-customers (citizens and businesses). In order to focus development efforts, a number of priority services for users have been identified for inclusion in digitalization initiatives. These services will be provided through a unified access point even when they involve more than one government department. In other words, the complexity of the public administration will not be apparent to users.

- **Digital identification** – Techniques for user identification and secure signatures adopting the Electronic ID Card, the National Services Card and Digital Suggestion.

- **Access channels** – a multiplicity of innovative channels for accessing services: the Internet, call centres, cell phones, third-party networks, etc.

- **Service provision agencies** – efficient and low-cost back office operations for service providers.

- **Interoperability and cooperation** – establishment of standards for interfaces between departments that permit efficient and transparent communication with the outside world.

- **Communication infrastructure** – A communication infrastructure that links all government departments.
2.4 Examples of e-governance projects

National Informatics Centre [143] has conceptualized, developed and implemented a very large number of projects for various central and state government ministries, departments and organizations. Many of these projects are continuing projects being carried out by various divisions of NIC at New Delhi Headquarters and State/District centers throughout the country.

Here are some of the most noteworthy projects to offer the citizens a glimpse of the multifaceted, diverse activities of NIC, touching upon all spheres of e-governance and thereby influencing the lives of millions of citizens of India.

Nation and State wide projects are

- Agricultural Marketing Information Network (AGMARKNET)
- Central Passport System
- Community Information Centres (CICs)
- Computerised Rural Information Systems Project (CRISP)
- Court Information System (COURTIS)
- Department of Agriculture Network (DACNET)
- Examination Results Portal
- India Image
- Land Records Information System (LRIS)
- National Hazardous Waste Information System (NHWIS)
- Public Grievance Redress and Monitoring System (PGRAMS)
- Spatial Data Infrastructure (SDI)
• Training
• Video Conferencing

Centre for development of advanced computing [144] has taken major initiatives in this area of e-governance offering solutions and services as,

• Online management and monitoring system
• Decision support system based on data warehousing for multipurpose household data
• Hospital Information System
• Laboratory Automation for Central Power Research Institute
• Networking in institutions to enable enterprise-wide data exchange
• Consultancy for developing framework for information system planning
• Portfolio Management Solution
• Telephone Revenue Billing (Pune Telecom)

2.5 **Advance Reservation in Grids**

In this section, a brief description on some advance reservation projects or systems for job and resource management in grids are explained.

2.5.1 **Examples and Discussions**

**Maui Scheduler**

Maui Scheduler [83], which was originally developed by the Maui High Performance Center (MHPC), has evolved into a community project, and is
currently maintained by Cluster Resources, Inc. The Maui Scheduler is an advanced cluster scheduler that supports advance reservation, optimization, job accounting and Quality of Service (QoS) policies, such as job prioritization, job preemption, and service access. The Maui Scheduler can act as a local resource manager where it has limited support for job queues and static resource partitioning to different users, groups or jobs. It can also support integration with other local resource managers, such as PBS Pro [94] [10] and Sun Grid Engine (SGE) [114], and collaboration with grid schedulers to access resource information, job staging facilities, and advance reservations.

For the Maui Scheduler, each reservation has three major components: a set of resources, a timeframe denoting starting and ending time, and an access control list [83]. To reserve the resources, a user needs to write a task description which contains the exact required number of attributes, such as processing elements, memory, and hard disk. Then, the Maui Scheduler will find available resources based on the given task description. To improve utilization, the Maui Scheduler uses backfilling methods, which execute smaller jobs waiting later in a queue, provided that they do not affect the start time of existing reservations. The Moab Workload Manager [84] which is a commercial version of the Maui Scheduler provides the same reservation features. However, it has other advanced functionalities, such as dynamic partitioning, user statistics, fault tolerance and integration with Globus [48].
Dynamic Soft Real-Time (DSRT) Scheduling System

Dynamic Soft Real-Time (DSRT) scheduling system [65] [66], developed by University of Illinois at Urbana-Champaign (USA), is a reservation-based CPU management system for soft real-time (SRT) applications. SRT applications, such as in multimedia, have soft deadlines or require a minimum guarantee QoS. Thus, they are tolerable towards minor delays or lower frame rates.

In the DSRT system resources are shared among the SRT applications. The CPU scheduler within the DSRT system is responsible for scheduling these tasks according to their reservation parameters and usage patterns. Thus, it has various scheduling mechanisms, such as Periodic Constant Processing Time (PCPT), Periodic Variable Processing Time, Aperiodic Constant Processing Utilization for maximum resource requirement, sustainable resource requirement, and constant resource utilization, respectively [42]. In addition, the CPU scheduler partitions the resources to allow other non-reserved or time sharing processes to be run in parallel. However, these tasks are to be executed by the local operating system.

The CPU broker of the DSRT system is responsible for administering reservation requests, and performing admission tests to find out resource availability by interacting with the CPU scheduler. In addition, the CPU broker negotiates with users by providing a list of alternative offers if the original request is rejected. Finally, the CPU broker allows the users to specify what to expect in case their reservations finish early or late. In case of the reservation finishes early, the user can choose between termination and scheduling another process. In case
the reservation finishes late, the user can choose whether to allow the CPU broker to preempt or extend it for a certain period of time.

The QoS-aware Resource Management System [65] is an extended version of the DSRT system that reserves additional resource types, such as network and memory. Each resource type is associated with a broker and a scheduler. Thus, the SRT applications need to negotiate with different brokers individually, or they can delegate this task to the QoS broker for simplicity.

**PBS Pro**

Portable Batch System, Professional Edition (PBS Pro) [94] [10], is a local resource manager that supports scheduling of batch jobs. It is the commercial version of PBS with added features such as advance reservation, security, cycle harvesting of idle workstations, information management, and automatic input/output file staging. PBS Pro can be installed on Unix/Linux and Microsoft Windows operating systems.

PBS Pro consists of two major component types: user-level commands and system daemons or services [10]. Commands, such as submit, monitor and delete jobs, can be first submitted through a command-line interface or a graphical user interface. These commands are then processed by the Job Server service. These jobs are eventually executed by the Job Executor service or MOM. In addition, PBS Pro enables these jobs to be submitted to Globus [48] via the Globus MOM service. Finally, the Job Scheduler service enforces site policies for each job, such
as job prioritization, job distribution or load balancing, and preemption. By default, the Job Scheduler uses the First In First Out (FIFO) approach to prioritize jobs, however, it can also use Round Robin approach, where jobs are ordered based on the group's usage history and resource partitions.

Reservations are treated as jobs with the highest priority by the Job Scheduler service. Hence, reservation requests need to be checked for possible conflicts with currently running jobs and existing confirmed reservations, before they are being accepted. Requests that fail this check are denied by the Job Scheduler service.

**Sun Grid Engine (SGE)**

Sun Grid Engine (SGE) is an advanced resource management tool for distributed computing environments [114]. It is deployed in a cluster and/or campus Grid testbed, where resources can have multiple owners, but they can also belong to a single site and organization. SGE enables the submission, monitoring and control of user jobs through a command line interface or a graphical user interface via QMON. In addition, SGE supports check pointing, resource reservation, and Accounting and Reporting Console through a web browser.

In SGE, resources need to be registered or classified into four types of hosts. The master host controls the overall resource management activities, and runs the job scheduler. The execution host executes jobs, while the submit host is used for submitting and controlling batch jobs. Finally, the administration host is
given to other hosts, apart from the master host, to perform administrative duties. By default, the master host also acts as an administration host and a submit host.

To manage resource reservations, each job is associated with a usage policy or priority, the user group, waiting time, and resource sharing entitlements [114]. Thus, the earliest available nodes will be reserved for pending jobs with higher priority by the SGE scheduler automatically. This reservation scenario is mainly needed to avoid the job starvation problem for large jobs. On the other hand, SGE can leverage an external scheduler, such as Maui Scheduler [83] to provide more comprehensive reservation functionalities.

**Globus Architecture for Reservation and Allocation (GARA)**

Globus Architecture for Reservation and Allocation (GARA) extends the Globus resource management architecture [48], by providing advance reservations and end-to-end QoS management for heterogeneous resources, such as compute nodes, storage elements, network bandwidth or a combination of any of these [52]. GARA uses Globus toolkit's information service for resource discovery, such as obtaining site-specific policies, system characteristics, and its current state.

GARA adopts a layered structure, where a Local Resource Allocation Manager (LRAM) provides reservation services specific to each individual resource type and a higher-level GARA External Interface handles issues, such as registration, resource discovery, and authentication of incoming requests. To handle bandwidth reservations or network QoS, GARA uses differentiated
service mechanisms (proposed by Blake et al. [15]) by implementing an expedited forwarding per-hop behavior, configuring the ingress routers that it controls, and deploying online admission control mechanisms to enable adaptive management of reservations [53]. To reserve compute nodes, GARA adopts the Dynamic Soft Real-Time (DSRT) scheduler [65] for real-time scheduling of tasks. Finally, to reserve storage elements, GARA interacts with Distributed-Parallel Storage System (DPSS) [9] to achieve high-performance data handling.

Any co-reservation or co-allocation agents can interact with GARA seamlessly, by implementing the required advance reservation and information service API or by using the Java CoG Kit package [37]. With these approaches, agents can find available resources, make the required reservations according to QoS, and submit jobs on behalf of applications or users.

**Highly-Available Resource Co-Allocator (HARC)**

Highly-Available Resource Co-Allocator (HARC) [64], developed by the Center of Computation & Technology at Louisiana State University (USA), is an open-source system for managing multiple reservations of various resources. This can be done by users sending reservation requests to HARC via its Java API or a command-line interface. Then, the requests are managed by HARC Acceptors. These Acceptors are responsible for interacting with an individual Resource Manager of a specific type, similar to GARA’s LRAM. Next, the Resource Manager communicates with a local scheduler to determine the resource
availability in the future for a particular request. Finally, the Resource Manager sends a message to users via Acceptors, whether it accepts or rejects the given reservation request. If the request is accepted, then it needs to be committed afterwards [64].

From the above description, HARC employs a two-phase commit protocol. To ensure the reliability of Acceptors and to prevent any missing messages, HARC uses Paxos Commit [62], a transaction commit protocol, where it uses multiple Acceptors for the same user to communicate with Resource Managers. With this approach, each Resource Manager will send the same message to multiple Acceptors. If the head or lead Acceptor fails, then other Acceptors will take its place automatically. In HARC, new types of resource can be integrated easily by creating new Resource Managers. To reserve compute nodes, the HARC Compute Resource Manager works with a local batch scheduler that supports advance reservation, such as Maui Scheduler [83] or Moab Workload Manager [84]. To reserve network bandwidth, the HARC Network Resource Manager acts as a centralized scheduler that oversees the overall management of network traffic for the entire testbed [63].

**G-lambda Grid Scheduling System**

The grid scheduling system, developed as part of the G-lambda project, is a web service system that is able to allocate resources in advance [8]. The aim of the G-lambda project is to build a standard web service interface among resource
management systems in grid and network computing [39]. The grid scheduling system consists of two main components: the Grid Resource Scheduler (GRS) and the Network Resource Management System (NRM).

The GRS is developed using Globus Toolkit 4 [55], a Java implementation of Web Services Resource Framework (WSRF). It handles reservation requests from applications or Grid portals. To reserve compute nodes, the GRS interacts with Computing Resource Manager (CRM) on each site. To reserve network bandwidth, the GRS communicates with Network Resource Management System (NRM). The NRM provides optical paths on a GMPLS-controlled network infrastructure. GMPLS is a generalization of the MPLS architecture, where it supports multiple types of switching other than label switching, such as lambda and fibre (port) [22].

To satisfy the user's QoS requirements, the scheduling module inside the GRS interacts with the CRM and/or NRM to locate available reservation slots using a depth-first search scheme [8]. However, new scheduling techniques can be easily incorporated into the module without affecting the rest of the system.

**Grid Capacity Planning**

The Grid Capacity Planning system [82], developed by the University of Innsbruck (Austria), targets to provide users with reservations of grid resource through negotiations, co-allocations and pricing. The system has a 3-layered negotiation protocol, where the allocation layer deals with reservations on a
particular grid resource, the co-allocation layer performs a selection of available nodes from all resources based on user's QoS and optimization constraints (e.g. operating system and cost of reservations), and the negotiation layer communicates with the user about suitable reservation times and their prices. However, the system only concentrates on reserving compute nodes in advance. This is done by having the allocator and co-allocator components as WSRF web services based on the Globus Toolkit 4 [55].

The allocator exists at an individual grid site, where it uses a Vertical Split and Horizontal Shelf-Hanger (VSHSH) algorithm [82] to solve the allocation problem. In the VSHSH algorithm, nodes are dynamically partitioned into different shelves based on demands or needs. Each shelf is associated with a fixed time length, number of nodes and cost. A new reservation request is placed or offered into an adjacent shelf that is more suitable.

Then, the co-allocator collects results from allocators of various grid sites, and produces suitable reservation slots or offers according to the user's QoS requirements. To manage different requests from the user, the co-allocator delegates these tasks to co-allocation managers (CM). Each CM manages and negotiates one user request. Thus, this approach reduces the complexities in administering reservations.
Task Graph

With regards to the efficiency analysis of functional parallel programs, i.e. executing two or more tasks concurrently, there are only several works done so far. Sinnen and Sousa [91] analyze the efficiency of task graph schedules, such as Economical Critical Path Fast Duplication (ECPFD) [49], Dynamic Level Scheduling (DLS) [35] and Bubble Scheduling and Allocation (BSA) [134] with respect of different Communication-to-Computation (CCR) values. The authors report that the utilization of a resource drops down if the CCR value is increased, and it also depends on the network topology of the target system. Moreover, they find that for coarse grained parallel programs (low CCR), the efficiency achieved is lower than 50%. However, it can be easily shown that this definition of efficiency is equivalent to the earlier description.

Hoenig and Schiffmann [125] also compare the efficiency of several popular heuristics, such as Dynamic Level Scheduling (DLS) [35], Earliest Time First (ETF) [56], Highest Levels First with Estimated Times (HLFET) [117] and Modified Critical-Path (MCP) [74]. They use a comprehensive test bench that is comprised of 36,000 TGs with up to 250 nodes. Essentially, it reveals that the efficiency of these schedules is mostly below 60%, which means a lot of the provided computing power is wasted. The main reason is due to the constraints of the schedule as mentioned earlier. Therefore, the main goal of our work is to increase the efficiency of these TGs by interweaving them, and backfilling with other independent jobs (if applicable).
For running DAG applications in the cluster or Grid computing environment, there are some systems available, such as Condor [16], GrADS [28], Pegasus [23], Taverna [120] and ICENI [112]. However, only ICENI provide a reservation capability in its scheduler. In comparison to our work, the scheduler inside ICENI does not consider backfilling other independent jobs with the reserved DAG applications. Hence, the ICENI resource scheduler does not consider the efficiency of the reserved applications towards resource utilization. A comprehensive survey on the characteristics and functionalities of these systems and others is mentioned in [61].

2.6 Scheduling in Grids

In this section, a brief description on some scheduling projects or systems for job and resource management in grids is presented.

2.6.1 Examples and Discussions

Mariposa [75], which is a distributed resource management systems and the pricing based on load and historical information. It supports budget-based query processing and storage management. Mungi [36] was developed in the University of New South Wales and it provides the renting storage space that increases as available storage runs low, forcing users to release unneeded storage. It supports storage objects based on bank accounts from which rent is collected for the storage occupied by objects.
Popcorn [85], which was developed in Hebrew University. It gets access to resource and it transfers credits from buyer to the seller account. Popcorn parallel code runs within a browser of CPU cycles seller. Popcorn API-based parallel applications need to specify a budget for processing each of its modules.

Enhanced MOSIX [133] was developed in Hebrew University, Israel where the resource cost of each node is known. It represents the clusters of computers. It supports process migration such that overall cost of job execution is kept low.

Spawn [58] (Xerox PARC) Second-price/Vickery auction (uses sponsorship model for funding money to each task depending on some requirements) Network on workstations. Each WS executes a single task per time slice. It supports execution of concurrent program expressed in the form of hierarchy of processes that expand and shrink size depending on the resource cost.

CSAR Supercomputing center [58] (University of Manchester) Commodity market and priority-based model (they charge for CPU, memory, storage, and human support services) MPPs, Crays, and Clusters, and Storage servers. Any application can use this service and QoS is proportional to user priority and scheduling mechanisms.

Nimrod-G which was originally developed in Monash University. It supports economy models such as commodity market, spot market, and contract-net for price establishment. It is a real system that supports deadline and budget constrained algorithms for scheduling task-farming and data parallel applications.
on world-wide distributed resources depending on their cost, power, availability and users quality of service requirements.

GridSim [101], was developed in Monash University. Currently, it supports economic models similar to those used in Nimrod-G, but limited to them. A Java-based discrete event toolkit for simulating Grid resources, users, applications, and brokers. The economic Grid resource broker supports deadline and budget based time, cost, cost time, and conservative time optimisation scheduling algorithms.

**AppLeS: A Network Enabled Scheduler**

The AppLeS [27] (Application Level Scheduling) project at the University of California, San Diego primarily focuses on developing scheduling agents for individual applications on production computational grids. It uses the services of Network Weather Service (NWS) to monitor changes in performance of resources dynamically. AppLeS agents use static and dynamic application and system information while selecting a viable set of resources and resource configurations. It interacts with other resource management systems such as Globus, Legion, and NetSolve to implement application tasks. The applications have embedded AppLeS agents and thus become self-schedulable on the Grid. The concept of AppLeS has been applied to many application areas including Magneto hydrodynamics [43], Gene Sequence Comparison, and Tomography [107]. AppLeS can be considered to have a predictive heuristic state estimation model with online rescheduling and application oriented scheduling policies.
Condor: Cycle Stealing Technology for High Throughput Computing

Condor [73] [57] is a high-throughput computing environment developed at the University of Wisconsin at Madison, USA. It can manage a large collection of computers such as PCs, workstations, and clusters that are owned by different individuals. Although it is popularly known for harnessing idle computers CPU cycles (cycle stealing), it can be configured to share resources. The Condor environment follows a layered architecture and offers powerful and flexible resource management services for sequential and parallel applications. The Condor system pays special attention to the computer owner’s rights and allocates their resources to the Condor pool as per the usage conditions defined by resource owners. Through its unique remote system call capabilities, Condor preserves the job’s originating machine environment on the execution machine, even if the originating and execution machines do not share a common file system and/or user ID scheme. Condor jobs with a single process are automatically checkpointed and migrated between workstations as needed to ensure eventual completion. The Condor has been extended to support submission of jobs to resources Grid-enabled using Globus services [59].

Condor can be considered as a computational Grid with a flat organization. It uses an extensible schema with a hybrid namespace. It has no QoS support and the information store is a network directory that does not use X.500/LDAP technology. Resource discovery is achieved through centralized queries with periodic push dissemination. The scheduler is centralized.
Data Grid

CERN, the European Organization for Nuclear Research, and the High-Energy Physics (HEP) community have established an International Data Grid project [128] with intent to apply the work to other scientific communities such as Earth Observation and Bioinformatics. The project objectives are to establish a research network for data Grid technology development, demonstrate data Grid effectiveness through the large-scale real world deployment of end-to-end application experiments, and to demonstrate the ability to use low-cost commodity components to build, connect, and manage large general-purpose, data intensive computer clusters.

The data grid project focuses on the development of middleware services in order to enable a distributed analysis of physics data. The core middleware system is the Globus toolkit with extensions for data grids. Data in the order of several Petabytes will be distributed in a hierarchical fashion to multiple sites worldwide.

Globus: A Toolkit for Grid Computing

Globus [48] provides a software infrastructure that enables applications to view distributed heterogeneous computing resources as a single virtual machine. The Globus project is an American multi-institutional research effort that seeks to enable the construction of computational grids. Currently the Globus researchers are working together with the High-Energy Physics and the Climate Modeling community to build a data grid [4]. A central element of the Globus system is the
Globus Toolkit, which defines the basic services and capabilities required for constructing computational Grids. The toolkit consists of a set of components that implement basic services, such as security, resource location, resource management, data management, resource reservation, and communications. Globus is constructed as a layered architecture in which higher-level services can be developed using the lower level core services. Its emphasis on the hierarchical integration of grid components and their services. This feature encourages the usage of one or more lower level services in developing higher-level services. Globus services have been used in developing many global schedulers, including Nimrod-G, AppLeS, and Condor/G.

**Javelin**

Javelin [80] is a Java based infrastructure for internet-wide parallel computing. The three key components of Javelin system are the clients or applications, hosts, and brokers. A client is a process seeking computing resources, a host is a process offering computing resources, and a broker is a process that coordinates the allocation of computing resources. Javelin supports piecework and branch-and-bound models of computation. In the piecework model, adaptively parallel computations are decomposed into a set of sub computations.

The Javelin system can be considered a computational grid for high-throughput computing. It has a hierarchical machine organization where each broker manages a tree of hosts. Resources are simple fixed objects with a tree
namespace organization. The resources are simply the hosts that are attached to a broker.

Javelin follows a decentralized approach in scheduling, using work stealing and a fixed application oriented scheduling policy. Whenever a host completes an assigned job, it requests work from peers and thus load balancing is achieved.

**Legion: A Grid Operating System**

Legion [106] is an object-based metasystem or grid operating system developed at the University of Virginia. Legion provides the software infrastructure so that a system of heterogeneous, geographically distributed, high performance machines can seamlessly interact. Legion provides application users with a single, coherent, virtual machine. The Legion system is organized into classes and metaclasses.

Legion objects represent all components of the grid. Legion objects are defined and managed by their class object or metaclass. Class objects create new instances, schedule them for execution, activate or deactivate the object, and provide state information to client objects. Each object is an active process that responds to method invocations from other objects within the system. Objects can be deactivated and saved to persistent storage. An object is reactivated automatically when another object wants to communicate with them. Application level schedulers such as Nimrod-G [100] and AppLeS [43] can change Legion default scheduling policies with user-centric policies.
PUNCH: The Purdue University Network Computing Hubs

PUNCH [86] [87] is a middleware testbed that provides operating system services in a network-based computing environment. The PUNCH infrastructure allows seamless management of applications, data, and machines distributed across wide-area networks. Users can run applications via standard web browsers without requiring application changes. PUNCH employs a hierarchically distributed architecture with several layers. The layers interoperate with a distributed resource management system and a predictive performance modeling sub-system in order to make intelligent resource allocation decisions.

Each of the resource management systems presented are a single model for resource trading. They have been designed with a specific goal in mind either for CPU or storage management. In order to use some of these systems, applications have to be designed using their proprietary programming models, which is generally discouraging, as applications need to be specifically developed for executing on those systems. Typically, in a grid marketplace, the resource owners, and users can use any one or more of these models or even combinations of them in meeting their objectives. Both have their own expectations and strategies for being part of the grid.

Grid Resource Brokers (GRB) may invite bids from a number of Grid Service Providers (GSP) and select those that offer the lowest service costs and meet their deadline and budget requirements. Alternatively, GSPs may invite bids in an auction and offer services to the highest bidder as long as its objectives are
met. Both GSPs and GRBs have their own utility functions that must be satisfied and maximized.

Grid brokers (note that in a Grid environment each user has his/her own broker as his agent) may have different goals, and each broker tries to maximize its own good without concern for the global good. This needs to be taken into consideration in building automated negotiation infrastructure. In a cooperative distributed computing or problem-solving environment (like cluster computers), the system designers impose an interaction protocol (possible actions to take at different points) and a strategy (a mapping from one state to another and a way to use the protocol). This model aims for global efficiency as nodes cooperate towards a common goal. On the other hand, in Grid systems, brokers and GSPs are provided with an interaction protocol, but they choose their own private strategy, which cannot be imposed from outside. Therefore, the negotiation protocols need to be designed assuming a non-cooperative, strategic perspective. In this case, the main concern is what social outcomes follow given a protocol, which guarantees that each broker/GSP's desired local strategy is best for that broker/GSP and hence the broker/GSP will use it.

2.7 Summary

This chapter describes some recent works related to advance reservation and scheduling in grids. In grids, all the presented works are able to reserve and manage compute nodes over hetero- or homogeneous systems. This can be done by interacting with a local resource manager at each site. Few grid systems, such
as GARA [52], HARC [64] and G-lambda [8], can also reserve network bandwidth.

All the aforementioned systems are expensive and time-consuming to build, operate and maintain. Thus, these exercises may not be feasible to some researchers and students. In this thesis, GridSim is chosen as the simulator. GridSim is an open source simulator that provides comprehensive features, such as advance reservation of compute nodes, resource failure, network QoS, and simulation of data grids. Hence, with GridSim, researchers and students can model various scenarios in networks and grids.