Revisions & Altercations in the Administration of Assets for Mobile Grid Computing

Aditya Vidyarthi
Jodhpur National University
Jodhpur, Rajasthan, India
adityavidyarthi5@gmail.com

A. S. Saxena
Gwalior, India

Abstract

This paper provides a discussion on the Mobile Grid concept, the revisions that this implies in the Grid technology and the altercations that have to be met in order to address the new computing paradigm, mainly focused on the aspect of resource administration, which is the heart of every Grid functionality. Grid computing has recently migrated from traditional high performance and distributed computing to pervasive and utility computing based on the advanced capabilities of the wireless networks and the lightweight, thin devices. This has as result the emergence of a new computing paradigm which is the Mobile Grid.

1. Introduction

One of the most important things for understanding and grasping Mobile Grid computing is to have a unfailing and exact definition, or at least fortitude of what a Mobile Grid is. There are many attempts for the accurate definition of the Grid. How- ever the various approaches that have been made address in a high degree of accuracy the term Grid. The Grid can be viewed as a distributed, high performance computing and data handling infrastructure, that incorporates geographically and organization- ally dispersed, heterogeneous assets (computing systems, storage systems, instruments and other real-time data sources, human collaborators, communication systems) and provides common interfaces for all these assets, using standard, open, general-purpose protocols and interfaces [1]. However, it is also the basis and the enabling technology for pervasive and utility computing due to the ability of being open, highly heterogeneous and scalable.

The Mobile Computing is a general term relating the application of small, transferable, and wireless computing and communication devices. The devices includes Lap- tops with Wireless Local Area Networks technology, mobile phones, and Personal Digital Assistants (PDAs) with Bluetooth or Infrared Data Association (IrDA) interfaces.

The Mobile Computing centers on the prerequisite of providing access to information, communications and services everywhere, anytime and by any available means. The technical clarification for achieving this are not always easy to implement.

In fact, mobile computing requires the creation of communication infrastructures and the modification of computer networks, operating systems, and application programs [2]. The mobility issue implies some constraints that must be addressed since they limit the capability of a moving resource in contrast to a fixed one.

Mobile Grid, in relevance to both Grid and Mobile Computing, is a full inheritor of Grid with the additional feature of supporting mobile users and assets in a seamless, transparent, secure and efficient way. It has the ability to deploy underlying ad-hoc networks and provide a self-configuring Grid system of mobile assets (hosts and users) connected by wireless links and forming arbitrary and unpredictable topologies.

2. Mobile Grid as a prerequisite

Grid is already being productively used in many scientific applications where huge amounts of data have to be processed and/or stored. Such tough applications have created, justified and diffused the concept of Grid among the scientific community. As the amount of potential Grid users is really enormous, the accumulated data processing and storage requirements are at least comparable. Wireless devices (lap- tops and PDAs), with currently limited assets (low processing power, finite battery life and constrained storage space), would benefit from the opportunity of using a considerable amount of assets made available by all the other devices connected to the network [3]. In particular, mobile users might be the future users of this new technology. Moreover, we have nomadic users who travel and work only seldom at their offices [2]. Mobile Grid enables both the mobility of the users requesting access to a fixed Grid and the assets that are themselves part of the Grid. Both cases have their own limitations and constraints that should be handled [5]. In the first case the devices of the mobile users act as interfaces to the Grid enabling job submission, monitoring and administration of the activities in an 'anytime, anywhere' mode, while the Grid pro- vides them with a high reliability, performance and cost-efficiency. Physical limitations of the mobile devices make necessary the adaptation of the services that Grid can provide to the users' mobile devices.

In the second case of having mobile Grid re- sources, we should underline that the performances of current mobile devices are significantly increased. Laptops and PDAs can provide aggregated computational capability when gathered in hotspots, forming a Grid on site. This capability can ad- vantage the usage of Grid applications even in places where this would be imaginary. However, there will always be the question on why a Grid solution should be adopted in comparison to any other non-Grid Information Technology (IT) solution. Grid is not intended to be the 'panacea' to the problems related to IT domain. It is a promising emerging technology that has the ambition to provide more efficient and beneficial solution than its 'competitors' by enabling the simple 'connect and share' approach in the same manner as the current Internet search engines apply the 'connect and acquire information' concept. By this, Grids and mobile Grids can be the ideal solution for many large scale applications that are of active nature and require transparency for users. Grid will increase the job throughput and performance of the involved applications and will increase utilization rate of assets by applying efficient mechanisms for resource administration in the vast amount of its assets. It will enable advanced forms of cooperative work by allowing the seamless integration of assets, data, services and ontologies.

3. Definition of mobile Grid

It is a stage that should speak to the mobility problem by means of permitting both fixed and mobile users to have access to both fixed and mobile Grid assets utilizing transparently and efficiently the underlying technologies. Mobility involves a set of issues for both users and communication/collaboration mode that will be presented shortly in the follow-up. Roaming users move around between different terminals and ask session movement (user mobility). For example, a user can open a session on a workstation at his office and then ask its movement to a personal computer at home. Roaming users, instead, moving around with their own wireless devices, want to continue their sessions when they cross different localities, as it happens today in cell phones telecommunications (terminal mobility). In both cases the middleware should hide the details of the active utilization of service components, when the session movement is required, besides it should manage all the session hand-off transparently to the user and possibly to the application components. One more kind of mobility is the service mobility which is meant as the ability of the infrastructure to maintain the services for a user while he is mobile. These services should be independent from the devices and from the location of the user with respect to the network he is accessing. [9] Another aspect of mobility that should be considered is the distinction between wireless networking and mobile user interfaces. In wireless networking the problem relies in the fact that the connection is highly variable both in quality and in price, and moreover that disconnections happen unexpectedly, thus constituting synchronous interaction uncertain. In case of mobile user interfaces the altercations come from the restrictions imposed by the small device size. Small displays can only present limited amount of information, while absence of keyboard makes data entry hard [5].

It should be observed that mobility stresses the fact of the synchronous mode of communication. In a collaboration environment the most challenging aspect is to en- able a synchronous mode of co-operation

between the peer entities. While asynchronous mode is easier to be implemented in unreliable and complex environments, the synchronous mode introduces a set of constraints that are by nature difficult to handle. The shared objects are no more workspaces and databases where many users act on, but moreover mechanisms have to be applied for enabling shared applications and concurrent schemes for the session handling.

4. What will get altered?

The mobile Grid will introduce revisions to the general Grid concept. New functionalities of the Grid will be required since the old ones will not make use of all the capabilities that will be available. These functionalities will involve end-to-end solutions with emphasis on Quality of Service (QoS) and security, as well as interoperability issues between the diverse technologies involved. Enhanced security policies and approaches to address large scale and heterogeneous environments will be needed. Additionally, the volatile, mobile and poor networked environments have to be ad- dressed with adaptable QoS aspects which have to be contextualized with respect to users and their profiles.

Mobile Grids will make use of the value that many mobile users perceive due to the advanced capabilities of their mobile devices. These advanced capabilities refer to the comparison of today's mobile devices with the ones that existed in the past. Al- though mobile devices are subject to physical constraints due to their nature they have the ability of having computational and storage capabilities similar to PCs, high quality displays, multiple interfaces (for instance Bluetooth, Ethernet adapters, USB, Infrared) etc. This value can be converted into revenue for service providers.

This implies revisions in various business models and policy issues. Complex workflows for businesses will be needed and Virtual Organizations (VO) will be enriched with the opportunity for automatic federations and resource sharing schemes. Valuations of the diverse services and especially with respect to the Service/Sharing level agreement have to be adapted in relation to QoS aspects. Enterprises have to issue policies that will handle the conflict between public rights to their users and viable models for operation.

Moreover the fair use of Grid must be determined by reconciling rights of public access to assets and private ownership of infrastructure and assets.

In the continuation we present a short discussion on some of the most important altercations of Mobile Grids with respect to the resource administration topic. Of course many things will alter and raise their own altercations to be addressed in this new context. Our study will focus in resource administration topic which is a very critical subject for the efficient utilization of the Grid infrastructure functional entities.

5.Altercations of Resource Administration in Mobile Grids

5.1. Resource Unearthing and Choosing

The unstable and active environment of the mobile Grid makes necessary the use of hi-tech mechanisms for resource discovery and selection. The inventory of the authorized machines and assets that are available in the Grid infrastructure is continuously updated. The selection of those assets that are expected to meet the time or cost constraints imposed by the user has to be based not only on deterministic criteria but also on stochastic parameters, thus providing complex probabilistic models for the topology of the Grids and their capabilities. Some of the required parameters are resource accessibility, system workload, network performance, etc. Not to forget that the target audience of the mobile Grids is the broad spectrum of the citizens who would like to use the Grid for solving problems of their everyday life. Thus, a financial criterion of the assets used should be under consideration for the proper re- source selection.

5.2. Scheduling of Jobs

Scheduling of Jobs is classified as an NP-complete problem [6] [7] due to the fact that an algorithm has to be applied that would allocate jobs to assets in an efficient and cost effective manner, so as to minimize the resource utilization gaining the maximum profit and satisfying at the same time the user constraints (Security, Quality of Service, Fault tolerance etc). In a mobile Grid environment there are many more constraints that would make the job scheduling problem more complicated. The cost of assets is in that case a

metric that is subject to context parameters. Imagine what can happen in a big conference room where the aggregated performance due to the many notebooks linked in wireless hotspots can result in a very cost effective solution for "instant" high performance computing. So the optimization criteria for a job scheduling mechanism should take under consideration not only the cost and the performance of each resource, but also the current availability of assets in the context and their reliability in providing the specific assets for the execution of the job under the given QoS constraints.

5.3. Replication of job, Migration and its Monitoring

After agreeing to the earlier Job Scheduling problem, we claim that the mobile grid is an active system where environmental conditions are subject to unpredictable revisions: system or network failures, system performance degradation, addition of new machines, variations in the cost of assets, etc. In such context job migration and re-scheduling, as well as job replication and co-scheduling are both (from different perspectives) efficient ways to guarantee the completion of the jobs according to the restrictions set by the users. Job monitoring and check pointing (especially in the case of long running tasks) is difficult in active environments. Especially job monitoring is responsible for detecting alert situations that could initiate a migration and alternatively identify if a job has been completed so as to suspend/stop/cancel the execution of the other replicas.

5.4. Administration of Replica for huge Data Sets

Replica administration is an important issue for a number of applications that utilize huge data sets (like for instance experimental results for scientific applications, for simulations and prediction models) [8]. The complete data set of a VO may exist in one or possibly several physical locations but it is possible that no single organization has sufficient storage to hold a complete copy. Instead, the common practice is to have copies of the most relevant portions of the data set stored on local storage for faster access. Replica Administration keeps track of where portions of the data set can be found by providing mappings between logical names for files and one or more copies of the files on physical storage systems. This mapping is of particular interest in mobile environments, since access to specific physical locations might be impossible. Storage devices are constantly changing their connection status and sometimes their physical location, thus imposing different schemes for accessing their data, which includes context aware information. Moreover, mobile Grids have to address the issue of not only locating a Data Base containing the relative data, but also if this location is the most suitable and reliable (with respect to avoiding a failure) to be used.

6. Conclusions

In this paper we presented a discussion on Mobile Grid definition, the basic motivation and benefits that can branch from this new technology, especially as far as it concerns collaborative engineering, industry and the everyday life of the "mobile" citizen. In the sequel it presents a set of major technical topics that influence the traditionally basic Grid Infrastructure services under the new Grid environment involving Mobile Grids**.** This topic indicates the necessity for research in these domains so as to enable the implementation of the new Mobile Grid environment. Moreover, the paper presented existing work and projects on worldwide level which are promising for the elaboration of thorough study in the domain. As the Grid computing seems to gain more space in the technology areas, it will be the basis for utility and pervasive computing in the forthcoming years by leveraging the mobility advantage that it will be enhanced with. The future work in this area will investigate not only issues concerning the interoperability but mainly with respect to the efficiency of mobile Grid solutions so as to enable the deployment in commercial and everyday life applications.

7. References

[1] Foster, "What is the Grid? A Three Point Checklist", GRID Today, July 20, 2002
[2] S. Gai, "Internetworking IPv6 with Cisco Routers", avail. at http://www.ip6.com
[3] D. Bruneo, M. Scarpa, A. Zaia, A. Puliafito, "Communication Paradigms for Mobile Grid Users", 3rd International Symposium on Cluster Computing and the Grid May 2003, Tokyo, Japan
[4] M. Graf, "Grid collaboration in the mobile environment." IBM. Speech at the Collaboration@work Workshop on June 25th 2003.
[5] S.-M. Park, Y.-B. Ko, and J.-H. Kim, "Disconnected Operation Service in Mobile Grid Computing," International Conference on Service Oriented Computing(ICSOC'2003), Trento, Italy, Dec 2003

[6] R. Moreno, “Job scheduling and Resource Management Techniques in Dynamic Grid Environments” 1st European Across Grids Conference, February 2003, Spain
[7] A. Abraham, R. Buyya, B. Nath, “Nature’s Heuristics for Scheduling Jobs on Computational Grids”, International Conference on Advanced Computing and Communications.