CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. 2008; **20**:997–1024
Published online 16 July 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cpe.1249

InterGrid: a case for internetworking islands of Grids

Marcos Dias de Assunção, Rajkumar Buyya*,†
and Srikumar Venugopal

Grid Computing and Distributed Systems (GRIDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Parkville, Melbourne, Vic. 3010, Australia

SUMMARY

Over the last few years, several nations around the world have set up Grids to share resources such as computers, data, and instruments to enable collaborative science, engineering, and business applications. These Grids follow a restricted organizational model wherein a Virtual Organization (VO) is created for a specific collaboration and all interactions such as resource sharing are limited to within the VO. Therefore, dispersed Grid initiatives have led to the creation of disparate Grids with little or no interaction between them. In this paper, we propose a model that: (a) promotes interlinking of islands of Grids through peering arrangements to enable InterGrid resource sharing; (b) provides a scalable structure for Grids that allow them to interconnect with one another and grow in a sustainable way; (c) creates a global Cyberinfrastructure to support e-Science and e-Business applications. This work identifies and proposes architecture, mechanisms, and policies that allow the internetworking of Grids and allows Grids to grow in a similar manner as the Internet. We term the structure resulting from such internetworking between Grids as the *InterGrid*. The proposed InterGrid architecture is composed of InterGrid Gateways responsible for managing peering arrangements between Grids. We discuss the main components of the architecture and present a research agenda to enable the InterGrid vision. Copyright © 2007 John Wiley & Sons, Ltd.

Received 13 October 2006; Revised 7 May 2007; Accepted 8 May 2007

KEY WORDS: grid computing; internetworking; InterGrid; peering; decentralized resource management

*Correspondence to: Rajkumar Buyya, Grid Computing and Distributed Systems (GRIDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Parkville, Melbourne, Vic. 3010, Australia.
†E-mail: raj@csse.unimelb.edu.au

1. INTRODUCTION

The growing popularity of Internet-based communication, computing, storage, and software technologies has led to the emergence of the Grid computing paradigm, which allows secure and coordinated sharing of globally distributed resources. Grid computing supports a range of e-Science and e-Business applications [1–3], which are enabled by various technologies that have been developed over a decade. They include Grid middleware such as Globus [4], Legion [5], UNICORE [6], and gLite [7]; schedulers like Application Level Schedulers [8]; and resource brokers like Gridbus Resource Broker (GRB) [9], Nimrod/G [10], Condor-G [11], and GridWay [12]. Efforts have been made to achieve interoperability between different technologies by defining common interfaces and protocols relying on Web Services [13].

The ultimate goal of Grid computing is to enable the creation of a Cyberinfrastructure that allows scientists and practitioners to cope with the scale and complexity of both current and next-generation scientific challenges [14–16]. Toward this, various national programs have initiated e-Science projects to enable resource sharing and collaboration among scientists. Examples include Enabling Grids for E-science in Europe (EGEE) [17] and Open Science Grid [18]. There exist also several national and international initiatives such as TeraGrid [19,20], APACGrid in Australia [15], K*Grid in Korea [21], NAREGI in Japan [22], Garuda in India [23], E-Science Grid in the UK [24], and OurGrid in Brazil [25] among others [26,27] that aim to provide national Grid facilities.

Such Grids, however, follow a restricted organizational model wherein a Virtual Organization (VO) [28] is created for a specific collaboration and all interactions such as resource sharing are limited to within the VO. Therefore, dispersed Grid initiatives have led to the creation of disparate Grids with little or no interaction between them. We term these as 'Grid Islands' in order to denote the lack of connectivity and inter-operation. Figure 1 illustrates Grid islands around the world. Internetworking of islands of Grids is needed to provide a global Grid-based Cyberinfrastructure [29] that will enable adaptive applications that can be assembled by discovering available services, and grow and shrink in terms of resource consumption.

1.1. Internetworking Grids: motivations and challenges

We motivate the need for internetworking of Grids with an example. Scientists from U.S.A., France, New Zealand, and Australia have developed mathematical models of kidney functions and have been sharing these models via Grids [30]. It is easy to extrapolate this to sharing of different models related to other organs that are developed within Grids dedicated to them. In order to build a complete model of the human physiology (e.g. IUPS Physiome Project [31]) one needs the capability that supports composition of models from different Grids. These models may be discovered through distributed information services enabled by peering of such Grids, which are controlled by the respective communities. Realizing this scenario requires participants to allocate resources from different islands of Grids in a seamless manner, and permit peering among Grids, which are under different administrative policies and political boundaries. Drawing an analogy from the mature field of computer internetworking, Grid islands are similar to having isolated Wide Area Networks (WANs) where users are restricted to content within a network rather than the Internet wherein users can access content from anywhere in the globe. Enabling the next generation of adaptive Grid applications requires an evolution of Grid infrastructures much in the same way the

DOI: 10.1002/cpe

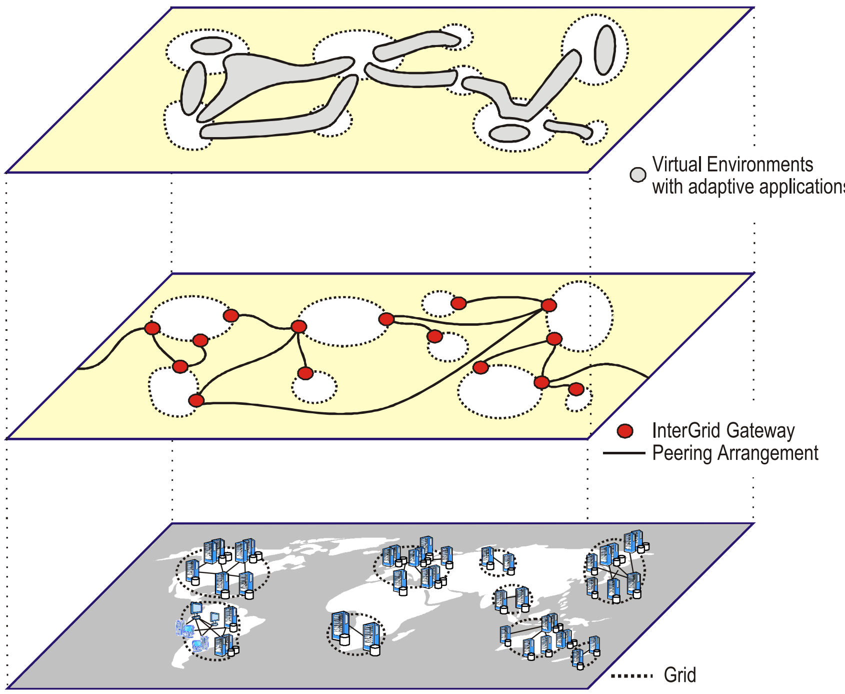

Figure 1. Internetworking of Grid islands to enable virtual environments and adaptive applications.

Internet evolved from isolated Local Area Networks to WANs and ultimately to a global network. We term the structure derived from internetworking Grid islands as the *InterGrid*.

In the Internet, ISPs establish arrangements, through which they agree to allow traffic into one another's networks. Such agreements are frequently termed peering and transit [32] and are enforced through routing policies. We envisage the InterGrid to be created in a similar fashion, by means of *InterGrid Gateways* (as shown in Figure 1) that mediate resource access based on peering arrangements between different Grids. Applications deployed on the InterGrid create *virtual execution environments* that are overlay networks whose topologies and resource consumption can vary over time based on the demands of the applications. These concepts will be explained in greater detail later in the paper.

The structure of current Grids does not follow principles such as the peering between Internet Service Providers (ISPs) present in the Internet [32]. While there are currently several efforts to promote interoperability between Grid facilities (e.g. 'Grid Interoperability Now' (GIN) at the Open Grid Forum (OGF)), they do not focus on peering arrangements between Grids. Interoperability and common protocols are important, but not enough to promote interlinking of Grid islands. For example, a set of common communication protocols underlies the Internet, but when ISPs peer with one another they consider their policies, economic issues, and the social and economic impact of peering. It is therefore important to identify the key issues of current Grid technologies that do not allow Grids to evolve to such a level. The questions that are to be answered are

DOI: 10.1002/cpe

as follows:

- What are the issues at the architectural level that prevent current architectures to scale to the InterGrid?
- What kind of architecture, peering arrangements and policies are required to realize the InterGrid?
- What are the coordination mechanisms that we need to put in place to enable the InterGrid?
- What are the incentives that would drive the end-users, laboratories, organizations, service providers, and Grid facilities to engage in such a network of Grids?
- What kind of agreements between Grids is needed for the InterGrid?

Therefore, there is a need for mechanisms and policies that support internetworking (also referred to here as interlinking or peering) of islands of Grids in a decentralized manner. Internetworking in this context refers to peering arrangements and mechanisms for InterGrid resource allocation, automated resource reservation, interconnection of information services and accounting, and cross-Grid scheduling.

Realizing the InterGrid requires Grids to adopt mechanisms that enable administrative separation by allowing network of networks, similar to many network-based systems such as the Internet, and numerous social and biological systems [33–36]. Such needs bring out various requirements for a comprehensive architecture, mechanisms and peering policies that allow: (a) Grid infrastructures to peer with one another and provide resources to one another when it is required; (b) provisioning and allocation of resources from one Grid infrastructure to the others; (c) policies and algorithms for InterGrid resource brokering; (d) application models that cope with the dynamics of peering of Grids; and (e) VOs and application environments to span multiple Grids.

1.2. Paper contributions and organization

This paper draws lessons from the growth of similar internetworked systems such as the Internet to envision the InterGrid as an evolvable system that can expand from organizational Grids to complex structures without major problems or scalability limitations. It, therefore, makes the following key contributions:

- Propose a model for the InterGrid as a system that allows the internetworking of islands of Grids to cope with a range of next-generation challenging tasks in business and scientific areas.
- Investigate successful global infrastructures such as the World Wide Web (WWW) and the Internet, and identify key aspects that influence their growth.
- Identify the weaknesses and shortcomings of the current Grid models that prevent the growth of the InterGrid.
- Develop an architecture for the InterGrid which allows peering between Grids, with possible policies and incentive mechanisms that can be used to ensure its sustainability.
- Propose a research agenda by identifying several challenges that need to be addressed, such as policy-based peering of Grids, pricing of Grid resources, coordination among Grids, enabling feasible market models, infrastructure for Grid economics, integration of accounting systems, automated resource reservation, InterGrid resource allocation, among others. Possible solutions to some of these challenges are also necessary to ensure the adoption of Grid computing within the industry.

The rest of the paper is organized as follows. Section 2 contains a description and analysis of the structure and principles that form the basis of the Internet and the WWW. Section 3 presents a gap analysis of existing Grid systems. We then present the structure of the InterGrid and the proposed architecture in Section 4. Section 5 then discusses the proposed research agenda on this topic. In Section 6 we present relevant work that can provide the technology required for enabling the InterGrid. Section 7 concludes the paper and presents our final considerations on the subject.

2. ANALOGOUS GLOBAL SYSTEMS AND THEIR PROPERTIES

As pointed out by Smarr [36], many infrastructures for current well-known services evolved from isolated initiatives that were later connected. In this section, we examine the existing infrastructures and draw some lessons from them for designing the InterGrid.

2.1. The Internet

The Internet has grown from a small project from DARPA started in 1969, initially linking a few sites in the U.S.A., to the millions of hosts and networks that currently comprise its intricate topology. The interconnection between ISPs is shown in Figure 2. The figure shows that hosts are

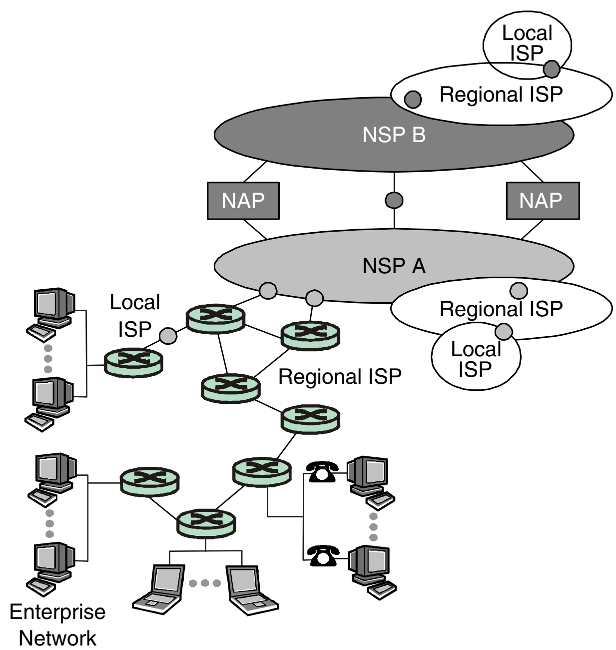

Figure 2. Interconnection of ISPs [37].

DOI: 10.1002/cpe

connected to local ISPs through access networks. In dial-up or broadband services, the local Public Switched Telephone Network loop is commonly used to provide users with access to the Internet. These local ISPs then connect to regional ISPs, which in turn, connect to national and international ISPs, also known as Tier-1 ISPs. Such national and international ISPs represent the highest level of the Internet hierarchy and are connected to each other either directly or through Network Access Points, also known as Internet Exchange (IX). Thus, the ISPs can provide services like access, backbone, content, application, and hosting. This structure has allowed the Internet topology to grow quickly and without the endorsement of a central authority [37].

Currently, the Internet presents an intricate structure comprised of a vast number of physical connections established by commercial contracts such as peering agreements [32]. Such agreements are legal contracts that specify the details of how ISPs exchange traffic. Norton [38] highlights the difference between peering and transit. Peering is the relationship whereby ISPs provide connectivity to one another's transit customers. Transit on the other hand, is the relationship through which an ISP provides access to all destinations in its routing table. The reasons for peering involve social, economical, and technological factors. ISPs can consider their policies, economical advantages, and conflicts before establishing agreements. Such agreements can be of various types, such as private, via IXs or in a relationship between customer and provider. They can specify the amount and proportion of traffic exchanged and the settlements since the traffic between peering ISPs can be asymmetric. Tier-1 providers, ISPs who have access to the global Internet, generally establish contracts not charging other Tier-1 providers, but they charge smaller ISPs for peering.

Another important concept is that of an Autonomous System (AS). In general, an AS comprises a network under a single administration and has its own policies to divert traffic or to avoid some peering ASs on the Internet. These policies are enabled by routing protocols such as Border Gateway Protocol, which allow ASs to advertise the routes that they prefer. An AS can have policies that take into account shortest or most cost-effective paths [39,40]. Such policy-based routing or peering is also applicable to the InterGrid, where Grids can favour a peering Grid more than others.

Some lessons can be learnt by analysing the structure of the Internet and how it has grown, such as:

- The Internet is a global network enabled by a common set of simple protocols that allow the interoperability among networks with different technologies, physical network interconnections, and varying peering arrangements.
- Even though the Internet has a complex topology, it has a structure that can grow quickly because there is no need of agreements and negotiations involving multiple organizations; a host in the Internet does not need to be directly connected to a large number of networks in order to have access to hosts in other networks.
- A self-healing structure, in which the failure of part of the network does not compromise the whole Internet.
- Although ISPs compete with one another, peering allows peered ISPs to provide global connectivity, reduce the amount of traffic across an expensive boundary and improve the efficiency for their customers [38,41]. In addition, its business model benefits end-users and compensates service providers [42].

DOI: 10.1002/cpe

- Routing protocols that allow traffic to be diverted when it is not allowed or viable to cross a specific network; these protocols allow ASs to deploy a range of routing policies based on internal interests.
- Networks are linked through routers in the Internet, therefore forming a large network of networks.

2.2. The World Wide Web and Content Delivery Networks

The WWW is one of the major network applications that contributed towards the rapid growth of the Internet [43]. Currently, the Web is a merge of network, protocols, and hypertext, which has led to the emergence of a plethora of scientific and commercial applications and several business models [44].

Although the WWW provides a global system on top of the Internet that allows the sharing of several kinds of media, there have been concerns about the performance and quality of the content delivered to Web users. Content Delivery Networks (CDNs) address some of these issues. A CDN is an infrastructure that replicates Web content from origin servers to replica servers (surrogates) placed in strategic locations. The main aim is to minimize internet traffic and response time by making clients retrieve files from nearby servers. Each CDN is set up and operated by providers such as Akamai [45] and Mirror Image [46]. In addition, several content providers have set up their own CDNs. A CDN can generally grow to a certain extent due to economical and technical reasons that prevent it from covering specific regions such as the high cost of over-provisioning [47]. Content Internetworking (CDI) [48,49] through the peering between CDNs has been proposed as a solution to this problem, allowing CDN providers to cover broader areas and minimize costs with infrastructures. However, CDI poses challenges like the definition of protocols and policies for internetworking of accounting systems, content distribution and request routing. The challenges imposed by CDI have some similarities with those of the internetworking of Grids. CDI, however, is a more mature concept. Therefore, we can draw lessons from such endeavours, such as the integration of the accounting systems, the decentralized allocation of resources, and the settlements among CDNs.

Some lessons to be drawn from the WWW and CDNs are as follows:

- Although the WWW relies on standard protocols, it has given rise to several business models that utilize the Internet infrastructure and the WWW in varying ways.
- The presence of network effect [50]. By providing an application for accessing hypertext and a variety of business models that rely on the Internet as network infrastructure, the WWW has provided incentives for users to join the network. This has resulted in the constant growth of the Internet and incentives for ISPs to remain interconnected.
- The limitations in speed and the latency of network links have led to the creation of CDNs. The replication and delivery of Web through CDNs is crucial in today's WWW. Many Web sites rely on CDNs to deliver increasing variety of content, such as hypertext files, images, videos, among others. The outsourcing and placement of services in strategic locations is also a reality in current WWW. It is important to note that CDNs have incentives for cooperating with one another, to alleviate flash crowd events, provide a better Quality of Service (QoS) to their customers and minimize the costs of expensive infrastructures [47].

3. LIMITATIONS OF CURRENT GRID SYSTEMS

When considering a large-scale system such as the InterGrid, a number of challenges arise, such as resource management among different Grids, varying resource usage across Grids, different security policies, resource reservation and co-allocation by research communities in peered Grids, agreements on QoS requirements and SLAs, and formation and management of VOs in the InterGrid. This section discusses some challenges that need to be addressed in order to realize the InterGrid vision.

Peering Arrangements: In the Internet, although there are standard protocols, ISPs have policies that define how the peering with other ISPs is performed. ISPs have agreements and implement the peering policies by defining what routes are preferable by considering the economic impact and incentives of peering with other ISPs. Work in Grid has focused on interoperability, but not on the peering between Grids and its economic implications.

Standards: A broad and well-attended standardization process has been going-on in the Grid community under the auspices of the OGF and based on the Open Grid Services Architecture (OGSA). The adoption of these standards is important for applications to be able to execute seamlessly over different Grids.

Different policies and mechanisms for resource allocation: Besides the interoperability between Grids at the middleware level, interlinking Grids requires advanced and automated mechanisms for InterGrid resource allocation, reservation, accounting, and scheduling. However, this is complicated by the different policies and mechanisms for resource allocation, followed within Grids, that may be incompatible with one another due to different levels of importance given to various resource usage criteria. This may create potential problems in reconciling different allocation policies or to create mechanisms to map policies from a Grid to another. An agreement on the standard criteria for resource usage and standard levels of QoS between federated infrastructures is therefore required for InterGrid resource allocation.

In Differentiated Services (DiffServ), Differentiated Service (DS) domains [51] can define Per Domain Behaviours (PDBs) [52] that have associated Per Hop Forwarding Behaviours [53], traffic classifiers, and conditioners that specify how a given traffic flow is treated within a DS domain. Particularly, a PDB aims to provide means to measure how a traffic flow is handled within the domain and is a building block for inter-domain QoS. A PDB specification intends to define under what conditions the output of a domain can be joined to another under the same traffic conditioning and expectations. PDBs can enable business agreements between ISPs regarding how one another's traffic flows are treated within their networks. The Grid community has been working towards enabling the recruitment of resources from multiple sites to process high-priority jobs. However, we envision that lessons can be drawn from other fields, such as DiffServ, to enable 'Per Grid Behaviours' and QoS guarantees across Grids.

Incentives for collaboration and attracting service providers: In the Internet, ISPs have incentives for cooperating and establishing peering agreements with one another. Consumers and enterprises have benefits in establishing their presence in the WWW. The InterGrid needs to provide incentives for equivalent participation of individual Grids and resource providers. A number of approaches have been proposed using economic models to address resource usage and incentives in a Grid [54–56]. Particularly, a well-designed market-based resource allocation mechanism provides incentives for participation by ensuring that all the actors in the system maximize their utility and do not have incentives to deviate from the designed protocol.

Pricing of resources and estimation of requirements: Markets can provide solutions for decentralized resource allocation in large networks such as the InterGrid. However, there are well-known concerns underlying their adoption by the distributed systems community, such as the pricing of resources. Several price setting mechanisms currently adopted by resource providers derive the prices for resources from auctions or utilize linear pricing functions. However, in several Grids, there is no mechanism for estimating the price of resources and no means for users to express the value of their applications in terms of currency. Thus, an important concern is how resources should be priced and how usage is measured. What would be the basic units of usage for a compute or a storage resource? How do resource providers adjust the price of their resources in a competitive Grid? What are the different possible price mechanisms in a Grid market, considering the local pricing, competitive market and collaboration among Grids? How do the price mechanisms affect the system? How do Grid users and organizations estimate their needs for resources? All these questions need to be answered before economic models can be applied to Grid computing successfully.

Connectivity and interaction patterns: The integration of Grids can enable a large number of interaction patterns, which would be difficult to design in terms of middleware, scheduling, and resource allocation. It is advocated that overlay networks will be important in a large-scale Grid to tackle this heterogeneity and guarantee several interaction patterns [57]. Overlay networks are virtual networks that cover physical infrastructures such as the Internet and add value to them with some features and semantics. They can enable various interaction models through Application Programming Interfaces to abstract the middleware from the complexity of the underlying network.

Coordination mechanisms: As demonstrated by Ranjan *et al.* [58], most of the current approaches to resource allocation are non-coordinated. Such approaches can lead to inefficient schedules and sub-optimal resource utilization. Coordination mechanisms that allow brokers and resource management systems to exchange information need to be put in place. However, the main challenge is that the InterGrid has Grids with different connectivity patterns. Thus, questions to be answered here are: what metaphors should coordination mechanisms follow? How can current mechanisms be improved to satisfy the InterGrid's requirements?

3.1. Limitations in Virtual Organizations (VOs)

Grid computing is also defined as the coordinated resource sharing and problem solving in dynamic and multi-institutional VOs [28]. A VO can be composed of a group of individuals and/or institutions that come together to share resources with a common purpose. According to Anastasiou *et al.* [59], the life cycle of a VO can be divided into: (a) the identification of a business opportunity; (b) the formation of a VO; (c) its operation and management; and (d) its termination. However, some problems can arise when considering these steps.

Formation of VOs: Currently, organizations define the terms for formation of VOs through multilateral contracts and agreements through offline processes. It is not possible to create VOs in an on-demand and dynamic manner due to security and policy-related issues. There is also a lack of mechanisms for the negotiation and establishment of agreements to dynamically form VOs. Moreover, a framework to define the off-line agreements for composing the source network or physical infrastructure is required. In addition, some legal barriers for the formation of VOs exist; some nations impose restrictions and require detailed information on the nature of collaboration with

DOI: 10.1002/cpe

scientists from other countries. There is, thus, a requirement for change in laws and legal processes for the establishment of VOs.

Resource management in VOs and across VOs: Providing a fair resource allocation in VOs is troublesome since resource providers can subscribe to multiple VOs and provide different amounts of resources to different VOs. Meta schedulers [60], some taking into account VOs, have been proposed. Dumitrescu *et al.* [61] highlight that challenging usage policies can arise in VOs that comprise participants and resources from different physical organizations. Participants want to delegate access to their resources to a VO, while maintaining such resources under the control of local usage policies. In this context, Dumitrescu *et al.* seek to address issues regarding the enforcement of usage policies at the resource and VO levels; mechanisms used by a VO to ensure policy enforcement; the distribution of policies to the enforcement points; and how to make job and data planners aware of the VO policies. Dumitrescu *et al.* [61] have proposed a policy management model in which participants can specify the maximum percentage of resources delegated to a VO. A VO in turn can specify the maximum percentage of resource usage it wishes to delegate to a given group of the VO. However, such policies are defined in an off-line basis and are complex to reconcile.

We believe that resource allocation in static and dynamic VOs could use the metaphor of a corporation. Shareholders that hold the most shares have the right to take decisions regarding how resources are allocated in the VO. The decision taker is chosen in the formation of the VO or as the VO evolves. However, it is important to have an accounting and ethic committee to avoid abuse in the VO. In addition to these problems, VOs can be recursive. That is, a VO can be composed of multiple sub-VOs. Resource allocation among these VOs has to account for the problems previously described, in addition to the allocation problems across these VOs.

Security in VOs: At present, Grid Security Infrastructure provides the basis for security in the Grid. At the VO level, VOMS [62] offers support to manage users, groups, roles, and capabilities in VOs. They allow a centralized control of VOs and extend Grid security concepts to a VO level by proving additional services which include: (a) VOMS server that maintains information about users, groups they belong to, roles and permissions; (b) a client that allows the user to create a VOMS proxy certificate; and (c) a VOMS administration service that allows the manager of the VO to set up roles and capabilities. There is also ongoing work on automated generation and negotiation of access control policies in VOs [63]. Issues regarding the mapping of existing privileges from a source domain to a target domain have also been investigated [64]. However, efforts are still necessary in this area in order to make the Grid a robust infrastructure for commercial applications that work within short lived, dynamic virtual execution environments.

In the next section, we present the InterGrid architecture that is designed to overcome the above limitations in Grids and promote the establishment of peering arrangements between them.

4. THE INTERGRID ARCHITECTURE

4.1. Network of networks structure

Through the investigation of existing infrastructures, we note that the concept of network of networks presented by the Internet is missing in Grid computing. In addition, the Internet aims at simplicity and providing a common set of protocols; the Grid is becoming a very complex architecture. Self-healing and benefits from peering, such as reducing traffic, increasing revenues or using services, are reasons adopted by ISPs for peering with one another. From the Web, we can see that the lack of centralized control has allowed for its fast growth. In addition, the Web has enabled a range of business models and organizations have reasons for using it. CDNs can peer to cover a broader area and share costs for an expensive infrastructure thus avoiding over-provisioning of resources and minimizing cost [47].

Based on communities and groups in our society and how they have formed, we see that such structures evolve from locally organized structures to those that are more complex [36]. For example, a group of individuals has a common interest on a given activity. Leaders of this group can look for another similar group and may find it useful to interact with another. After the agreement to cooperate has been settled, interactions can take place, new links can be made, and existing ones can be broken. Some tools have helped people to form communities of interests.

From the different structures analysed, we can note the following characteristics and needs:

- Small structures are linked to more complex ones through some access point. Examples include the Internet where routers link networks; and groups in the human society where leaders start agreements or collaborations with other groups.
- In joining and forming communities, there are places where people publish not just their capabilities, but also their interests and needs.
- Mechanisms allow one to locate and connect with people or organizations that can fulfil their needs.

4.2. Architectural requirements

In order to provide mechanisms that allow Grids to coordinate through peering arrangements, the InterGrid needs to meet the following requirements.

Incentive-oriented peering arrangements: Although the InterGrid is comprised of Grids with competing organizations with different and maybe conflicting interests, it will need to provide mechanisms that provide incentives for Grids to peer with one another. Like the peering between ISPs in the Internet, Grids require incentives to peer in the InterGrid and need policies for peering and trading resources of one another.

Standards based: The adherence to standards, such as instantiations of OGSA like WSRF are important to assure interoperability. The standardization allows automated decisions and policies to be implemented regardless the middleware or tools utilized by different Grids [13].

Respect to administrative management and separation: Grids can be created for various purposes, be under different administration and have different resource usage policies. It is important to respect both the internal policies and provide a structure that allows Grids to interlink with one another whilst respecting the concept of organization.

Deployment of applications that require resources from multiple Grids: Users can deploy applications and scientific workflows that require resources beyond the capacity of their Grids, thus requiring InterGrid resource allocation. These applications may need resources from multiple Grids for other technical reasons. For instance, the deployment of a CDN over the InterGrid may require surrogate servers to be deployed in multiple Grids.

InterGrid policies and decentralized resource management: A Grid requires a means to specify its policies, defining which resources are available to other Grids under which circumstances, and make them available to other Grids. However, a Grid will retain ultimate control over its resources and to whom it wants to provide access. We need to implement peering agreements between Grids without the need for global control over resources. Decentralized approaches for resource allocation such as self-organizing economic models [65,66] are required by the InterGrid.

Resource allocation, *reservation and brokering across Grids*: Each Grid has gateways that translate and reply to resource allocation requests originated in local and peering Grids. A gateway, aware of the peering arrangements with other Grids, forwards the request to other Grids that are able to provide the required resources. We need policies and mechanisms for selecting Grids to peer with and for admission control when accepting requests originated in other Grids. In addition, the InterGrid requires coordination mechanisms between gateways and means for reconciling policies or mapping policies from one Grid to another.

4.3. The proposed architecture

Based on our study of Internet and similar network-based systems, we propose the architecture shown in Figure 3 for the InterGrid. The architecture is a hierarchy that has the InterGrid Gateways (IGGs) on top coordinating between the different Grids followed by the IntraGrid Resource Managers (IRM) taking care of resource allocation using the resource shares assigned by resource providers.

The InterGrid proposed in this work can enable peering of Grids and exchange of resources. An application can very well demand resources from different Grids and perform resource management of its own. However, the development of applications that deal with peering agreements between Grids can be complex. We propose that applications can have performance and environment isolation provided by Dynamic Virtual Environments (DVEs) that can be deployed on top of the

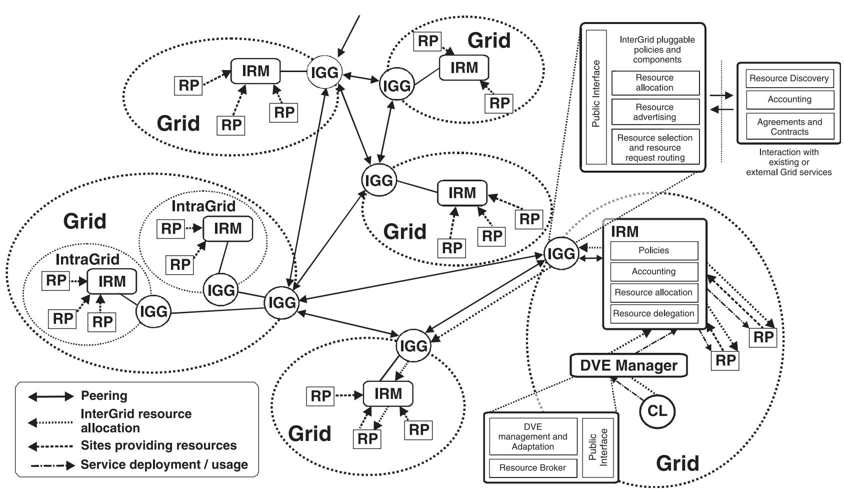

Figure 3. Architecture for internetworking of Grids.

DOI: 10.1002/cpe

InterGrid infrastructure [10,11]. DVEs provide consumers with a transparent network that behaves like a dedicated computing and data infrastructure, requiring little or no changes in existing Grid middleware and services. DVEs can span multiple Grids, and grow or shrink in terms of resource consumption [6] according to the demands of the applications running on them.

The proposed architecture, therefore, integrates the resource allocation represented by the InterGrid with the application deployment performed through DVEs. We describe the components of this architecture in detail as below.

Resource Providers (*RPs*): These contribute a share of computational resource, storage resource, networks, application services, or other types of resources to a Grid. The allocated share could be based on provisioning policies that are determined by the providers' perception of utility. For example, the owner of a cluster could allocate a share of the nodes in return for regular payments.

IntraGrid Resource Manager (*IRM*): The IRM manages the different shares of resources that have been allocated by the RPs to the Grid. The IRM is the first point of contact for acquiring resources for DVEs. Based on its policies, the IRM also assigns provisioning rights over a collection of Grid resources to an IGG. The IRM is kept functionally separate from IGG in the architecture because internally an individual Grid can have an internal resource management system already in place. In this case, the different kinds of resource allocation policies (determined inside a Grid and managed by the IRM) and peering arrangements (determined outside a Grid and managed by the IGG) can be reconciled.

InterGrid Gateway (*IGG*): An IGG is aware of agreements with other IGGs; acts as a Grid selector by selecting a suitable Grid able to provide the required resources; and replies to requests from other IGGs, considering its policies. IGGs with pluggable policies enable resource allocation across multiple Grids. An IGG is chosen and maintained by a Grid based on internal criteria. The IGG also interacts with other entities including Grid Information Services (GISs), resource discovery networks, accounting systems, and resource managers within peered Grids. A GIS provides details about the available resources; and accounting systems provide information on shares consumed by peering Grids.

The key functionality of the IGG is described as given below:

- *Resource selection*: The resource selection module carries out selection of resources from peering Grid infrastructures, according to agreements between Grids, or based on resource selection policies defined by the peered Grids. The IGG selects, negotiates with, acquires and ranks resources from peering Grids.
- *Resource allocation*: This component exposes the resources from a Grid to other peering Grids based on the provisioning policies. In addition, based on the provisioning policies, accounting information and monitoring services, this component makes decisions on the requests from peering Grids that should be accepted or refused.
- *Pluggable peering policies*: These policies underlie resource selection and allocation among Grid infrastructures. Economy-inspired policies can be created and plugged into the IGGs.

Distributed Virtual Environment Manager (*DVE Manager*): We suggest that a client application (indicated as CL in Figure 3) can acquire resources from the InterGrid by requesting the DVE Manager for the instantiation of a DVE. The DVE Manager, on behalf of the client, interacts with the components responsible for peering of Grids to acquire resources. The DVE Manager also monitors the resources that join or leave the DVE, deploys the services required by the client, and

DOI: 10.1002/cpe

drives the adaptation of the resource allocations based on the demands of the client. DVEs can accommodate existing Grid and virtualization technologies by enabling dynamic overlay networks on top of the InterGrid, whose topologies and allocations may change over time. Existing Grid technology can be deployed on the resources acquired for a DVE and can provide the information required by the DVE Manager to adapt the topology of a DVE or to change its allocations.

Client Application (*CL*): The client corresponds to an application that uses resources from the interlinked Grids. A VO can also be treated as a client that requires a DVE. In this case, a VO manager specifies the resources necessary to form the VO or increase the allocations of an existing VO, and the required services and Grid middleware that have to be deployed on the DVE. Once the DVE is created, members of the VO can utilize the resources.

In addition, the following components are required even though they are not included in the architecture.

InterGrid Directories and Marketplaces: An InterGrid Directory is a repository with information regarding Grids, Grid projects, their goals and capabilities, proposals for collaboration and requirements by Grid projects. The current facilitators for VOs such as OSG and EGEE could maintain InterGrid Directories with information that can be shared, such as existing VOs and their Grid projects.

4.4. Resource allocation model

We illustrate in this section how a client can obtain resources of the InterGrid for an application. The application can implement its own resource management mechanism or utilize existing Grid resource brokers. However, we envision that the application can request a DVE Manager to acquire resources from the InterGrid, create a DVE, deploy the required services, and manage resources comprising the DVE. The abstraction of containers is used for resource allocation across Grids. We use the term ‘slot’ to denote a Virtual Machine (VM) or a physical resource that can be acquired by the DVE Manager, on which the services and applications can be deployed. The workflow for resource allocation in the proposed architecture is shown in Figure 4 and can be described as follows:

1. A RP advertises its slots in the registry provided by the IRM. The advertisement is made through a ‘slot assertion’, which is a delegation of the provisioning rights over a set of slots or resource shares.
2. An IRM can delegate the provisioning rights over the resources within a Grid to the IGG by providing a part of or all the resource assertions given by RPs.
3. A client, who can be an application or a VO manager, shows interest in obtaining a number of slots to deploy an application. The client contacts the DVE manager, to which it provides a description of the required slots and the services that have to be deployed on the allocated slots.
4. The DVE Manager will: (i) acquire the resources; (ii) deploy the services; and (iii) manage the DVE that will be composed by the set of allocated slots. The DVE manager tries to acquire the slots needed by issuing ‘slot requests’ to the IRM.
5. The IRM can provide all or part of the required slots based on the provisioning policies. If the individual Grid is not able to provide the required slots to the DVE Manager, for performance or technical reasons, then, the IRM forwards part of the slot requests to the IGG.

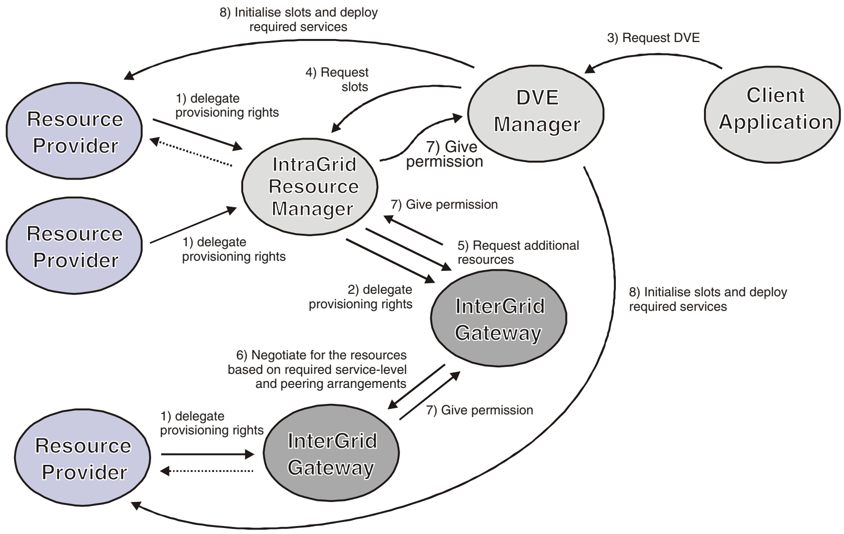

Figure 4. Resource allocation model for the proposed architecture.

6. The IGG, based on the peering agreements and the peering policies, selects the best-peering Grid from which the slots are allocated.
7. Once the slots are acquired, the DVE is given a permission to use them.
8. At the desired time, the DVE Manager will invoke the initialization of the resources at the RPs, passing the permission obtained beforehand. The DVE manager receives a permission that contains the description of the share, the slots obtained, and the duration time of the permission. The DVE manager, on behalf of the client, is responsible for contacting the obtained slots and deploying the services.

Once the resources have been allocated and the services have been deployed, the users of the client application can make use of the services and applications running on the DVE.

4.5. The InterGrid Gateway control flow

The IGG is the enabler for the InterGrid, being a key component for managing the agreements between Grids and enabling coordinated resource allocation across Grids. Therefore, we present the control flow within the IGG in greater detail in this section.

The internal architecture of IGG is shown in Figure 5. The control flow of this component is described as follows. IRMs offer resources to a Grid by registering Slot Assertions (SAs) with the IGG describing the characteristics of the resources made available to the InterGrid (1). The advertised SAs are stored by the IGG in a Grid Resource Inventory (2). The IGG is aware of peering arrangements with other Grids. It receives slot requests and replies to them by providing a list of the allocated slots. It can use the peering arrangements to acquire slots from other Grids. A slot request can be made by a DVE to an IRM, which can be forwarded by the IRM to the IGG

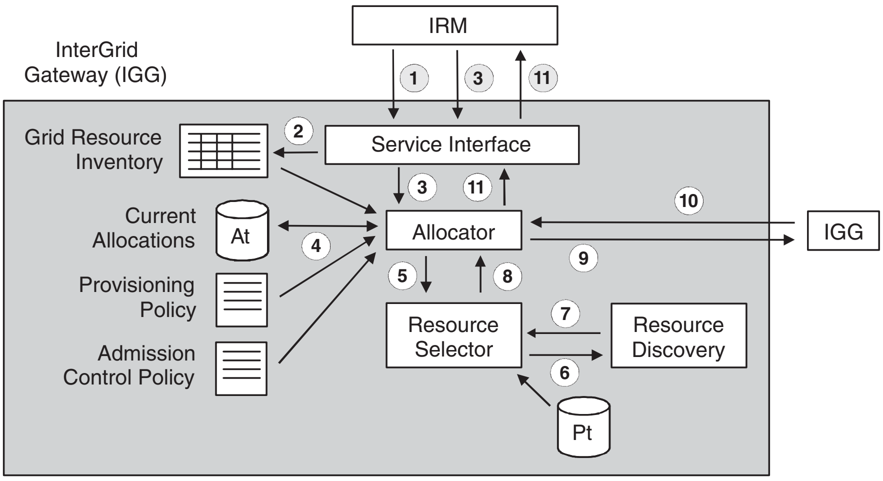

Figure 5. The IGG architecture: (1) IRMs delegate slot assertions (SAs) to IGG; (2) SAs are kept in the Grid Resource Inventory; (3) when a DVE within the Grid requires more slots than the IRM can provide, the slot request is passed to the IGG; (4) the policies and current allocations guide the Allocator in deciding, which slots can be allocated; (5) if there are not enough slots within the Grid, Allocator contacts Resource Selector to choose a Grid able to provide the required slots; (6) Resource Selector uses Resource Discovery to find a Grid that can provide the required slots. The peering policies, contracts and prices (Pt) are used to guide the decision on which peering Grid to choose; (7) the previous process returns a list of candidates to Resource Selector; (8) the candidates are passed to the Allocator; (9) Allocator sends the slot request to the selected peering Grid; (10) the IGG of the peering Grid sends a list containing references to the slots provided; and (11) the list of slots is passed to the IRM.

when the former is not able to provide the required slots (3). The slot request is then passed to the Allocator module.

The Allocator takes into account the availability of slots in the inventory, the provisioning and admission control policies, and the current allocations (At) when taking decision on what slots have to be allocated to the DVE (4). If the current Grid does not have enough slots to provide, part or the whole slot request can be passed to the Resource Selector component, which searches for a peering Grid that can provide the slots (5). The description of the peering arrangements with other Grids, policies applied and contracts are stored in a repository or maintained by a contract management system represented here by (Pt). Resource Selector utilizes this information as well as resource information obtained from resource discovery networks established between the Grids to determine what peering Grid can provide the required slots (6). The resource discovery network—or Resource Discovery—returns a list of candidates containing the peering Grids that can provide the slots (7). Once a potential peering Grid is found, Resource Selector reports it to the Allocator (8). The Allocator can then pass the slot request to the selected IGG (9). The selected IGG can return a list containing a description and references to the slots it is willing to provide or a reject message if the Grid does not have the required slots (10). The Allocator will inform the IRM about the overall list of slots allocated (11).

4.6. Mapping of the architecture to existing grid and virtualization technologies

The proposed architecture and mechanisms aim at leveraging existing Grid technologies. In this section, we discuss on how the proposed architecture maps to existing Grid and virtualization technologies. The advent of server virtualization technologies enables increasing flexibility in resource management as it enables performance isolation, migration, suspension, and resumption of virtual machines. Container-based systems allow the automated deployment of service applications. We propose that the resource control between Grids can be performed using the abstraction of containers as demonstrated by Ramakrishnan *et al.* [67]. Grids in such a scenario exchange resources on which services can be deployed based on their peering policies.

By using VM technology, the allocation of resources to a DVE resembles the creation of a slice in PlanetLab [68]. However, each IGG has some provisioning rights over the resources of its Grid. The architecture proposed shares some principles with the Global Environment for Network Innovations (GENI) [69] such as the use of entities with limited rights over resources of a country or region.

The creation of a DVE consists of the allocation of resources that can be VMs [70] and the initialization of these VMs. Existing technology that allows the deployment and scheduling of VMs can be used for the instantiation of these VMs; examples include Virtual Workspaces [71], Shirako [72], VIOLIN [73], In-VIGO [74] and Virtuoso [75]. Overlay network technologies can also be used to provide isolation and the feeling of a dedicated network infrastructure. The adaptation of DVEs proposed in this work can make use of principles such as migration of VMs, adaptation of network topologies [76], and autonomic deployment and migration of services.

Examples of technologies that can be used to provide the features required for a DVE Manager include the Service Manager component of Shirako and Grid resource brokers such as GRB [9]. GRB can be adapted for instantiating VMs or allocating resources on demand. An IRM can be a VM-based resource management system such as Shirako [72], VIOLIN [73], and Virtuoso [75].

5. RESEARCH CHALLENGES FOR THE INTERGRID

As the internetworking of Grid islands requires that fundamental issues be investigated and solved by the research community, we propose a research agenda to guide these efforts. The following agenda is defined based on the issues that have been identified and the lessons learnt from our analysis of analogous global systems and their principles. This agenda aims at grouping and defining some of the research topics that, in our opinion, deserve special attention in order to enable the vision of an InterGrid.

Grid computing has been moving towards supporting commercial applications in a manner similar to the Internet. In the Internet, ISPs are in the business to make profit—they see one another as competitors or sources of revenue—whereas users are interested in using services at low price. It has been shown that new design principles to the Internet have to accommodate such economic aspects [77]. Similarly, resource providers and consumers in different Grid islands will have different interests, which may be adverse to one another. These parties will act to favour their own interests and will share resources expecting financial compensation. Economy-based models are relevant to the InterGrid because resource allocation can be achieved through the economic behaviour of the involved parties, markets can provide incentives for providers to offer their resources to the

DOI: 10.1002/cpe

InterGrid, and can provide the settlements necessary between Grids. Therefore, in this research agenda, we focus on economic-based policies and mechanisms for interlinking Grids.

5.1. Formation, adaptation, coordination mechanisms and self-organization for the InterGrid

In this work, we advocate the need of IGGs as entities that are aware of agreements or peering arrangements among islands of Grids. These IGGs also select the best partners when there is a need to collaborate for the creation of a VO. In the InterGrid, islands of Grids with independent resource providers should cooperate to enable the co-allocation of resources and the formation of dynamic VOs. Thus, there is a need to specify how the whole system behaves and how IGGs can share information about current allocations and coordinate future allocations. This should preferably be performed in a self-organizing manner. Current coordination approaches neither define how the peering arrangements take place nor specify how the InterGrid will self-adapt. Thus, new methods and mechanisms are necessary to enable the peering of Grid islands, and the automated and responsive creation, operation, maintenance, and dissolution of VOs or DVEs in the InterGrid. We consider that a VO can span multiple Grids. Hence, once the peering arrangements are defined, it is important to predict how VOs will form, manage themselves, and ultimately dissolve in this environment. In addition, mechanisms to enable the formation and evolution of self-adaptable topologies for DVEs based on the requirements of the applications running on them are necessary. For example, data-intensive applications may require peers to be located closer or links to be built dynamically among the peers involved [78,79].

There is a need to solve complex tasks such as the cooperation between IGGs to enable the creation of a DVE with specific QoS requirements regarding issues such as resource availability and network topology. Moreover, it is important to adapt the allocation of resources to a DVE during its life cycle according to the demand of the applications running on it. A DVE can grow or shrink in terms of resource consumption and possibly change its topology. Thus, self-adapting and self-organizing resource allocation is necessary.

Considering these problems, some questions to be answered are:

- What are the mechanisms necessary to enable the formation and self-organization of overlay networks for the InterGrid?
- Using social mechanisms such as that proposed by Singh and Haahr [79], what are the models and techniques that could be used to promote the fast convergence and evolution of the topologies of these overlays on the InterGrid?
- What are the mechanisms necessary to enable the formation and operation of dynamic VOs in the InterGrid?
- What would be the growth patterns of the InterGrid based on chosen models and mechanisms?

Current resource allocation and scheduling mechanisms utilized by Grids are non-coordinated (i.e. different domains have their own resource brokers, schedulers, objectives, and QoS requirements and in many cases these entities do not exchange information about their allocation decisions). Such divergent approaches can lead to scenarios with bad schedules and inefficient resource allocation in the InterGrid. We envision that IGGs should coordinate allocation decisions through peering arrangements to enable the execution of applications spanning multiple Grids. We also support the need for self-organizing coordination mechanisms [66,80] for the InterGrid. In this case, the global allocation emerges from IGGs designed to interact with one another on behalf of the applications within the Grids that they represent without the requirement for global control or a centralized system [80,81].

We plan to apply computational economy as a metaphor for the internetworking of islands of Grids and formation of VOs in the InterGrid. Thus, in this scenario, some questions that need to be answered are:

- What type of market model is suitable for the InterGrid?
- What coordination mechanisms are required to allow the scalability of the InterGrid?
- What are the requirements and issues of such coordination mechanisms? Are there any protocols that can be used for exchanging information between IGGs?
- Is it possible to achieve adaptation and equilibrium through entities only engineered to achieve local self-organizing behaviour?
- What are the metaphors and models that can be used for self-organization in the InterGrid? How can they be applied in terms of design and development?

5.2. Peering agreements between Grids and related policies

Standard protocols ensure interoperability in the Internet and enable a host to send packets to any other host connected to the Internet. Although ISPs interconnect with one another, they place varying costs on routes and consider various criteria to carry out inter-ISP routing [32,39,40,82]. The two chief objectives guiding ISPs are: the incentives and viability of peering; and the minimization of costs by choosing one peering partner over another.

Principles similar to the policy-based routing in the Internet are applicable to Grid internetworking in activities including offloading and redirection of resource allocation requests to peering Grids. However, there are important differences. While Internet routing has to only consider data packets, peering of Grids is more complex as it involves managing resource allocation requests that could involve numerous attributes depending on the QoS expectations of the user. Therefore, the span of a DVE or a VO on a multi-Grid environment is determined by the complexity of its requests and the availability of resources that satisfy its requirements.

Resource allocation in a multi-Grid environment brings up interesting issues due to the dynamic environment and the varying demand for resources from applications within each Grid infrastructure. Efforts have been made to provide mechanisms for resource control and allocation of resources in a multi-provider site environment [72,83], but little has been done on mechanisms and policies for InterGrid resource allocation. In addition, mechanisms for increasing or releasing allocations are needed to adapt applications to a multi-Grid environment.

Grid infrastructures can establish contracts or Service Level Agreements (SLAs) that state the conditions under which they will peer with one another. A Grid infrastructure has its policies regarding how its resources will be allocated to the peering Grid. For example, Grid A can provide a best effort service for a peering Grid B, in which A tries its best to provide the resources required by B, when B needs to offload the demand for resources in its infrastructure. Grid A expects some financial compensation from B for the resources utilized. In another case, Grid A can have a different contract with Grid C in which A stipulates that it will provide 100 computing resources at maximum for no more than 60 hours per month. Grid A expects an equal compensation in terms of computing resources from C. Similar to the peering of ISPs [82], these agreements and settlements must be considered when an IGG requests additional resources from peering Grids to create or augment DVEs.

Relevant work has been carried out in understanding and defining policies for peering of ISPs [32,38], CDNs [48,49] and Peer-2-Peer networks [84]. The economics of peering are currently better understood in domains such as the Internet. In the Internet, ISPs are in the business to make profit—they see one another as competitors or sources of revenue—but they peer for economic or technical reasons [41,77,82,85]. Similarly, we expect Grids to peer based on economic incentives based on mechanisms such as financial compensation or bartering of peak loads. Thus, the investigation of peering agreements, mechanisms and settlements among IGGs is required. The specification of policies, policy-based selection and allocation of resources from peering Grids, and game-theoretic treatment of peering Grids are part of this research agenda.

5.3. Pricing resources in the Grid economy

The adoption of economic principles to Grids comes from observing the success of economic institutions in the real world as a sustainable model for regulating the allocation of resources, goods, and use of services [86]. However, the adoption of such economic approaches requires the study of pricing of Grid resources and/or agreement on pricing mechanisms. Therefore, if an economic approach is used by the InterGrid, detailed studies have to be done in areas such as resource pricing, modelling consumer's utility, resource provider's marginal cost, and benefit in providing resources.

Some of the questions that need to be answered are:

- What resources should be free of charge and what resources should be priced in the Grid market?
- What are the policies that define the circumstances for sharing resources in the Grid?
- How to price the resources in the Grid?
- What kinds of issues related to the price setting for the resources arise in the Grid?
- How do resource providers adjust the price of their resources in the Grid in order to achieve the price that maximizes their profits in a competitive market, yet maintaining the equilibrium of supply and demand?
- How do the price-setting mechanisms differ from one another when considering the local pricing, a competitive market and collaboration among Grids?
- How do different price mechanisms impact the system?
- How to model the resource price variation process to predict the future price of resources in the Grid?

Grid economy can become quite complex when the case of InterGrid is considered. Thus, the study of pricing of resources and its effect in the Grid economy may be studied according to the steps shown in Figure 6. Pricing of resources involves the following steps/issues:

1. *Cost/Benefit*: First, the pricing of resources within a local Grid is studied. The challenge here is how resource providers should calculate the cost and benefit of providing the resources, in order to determine their prices in the local Grid.

DOI: 10.1002/cpe

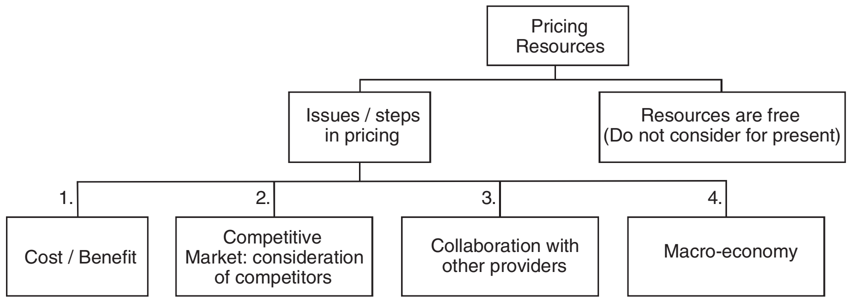

Figure 6. Pricing of resources steps/issues in the Grid.

2. *Competitive market*: Thus, resource providers must consider different ways of adjusting their price within a competitive market with the aim of maximizing their profits. The target price in the market should be the one that is at the equilibrium of supply and demand, and maximizes the overall welfare. In this regard, equilibrium theories are to be considered. Models to be developed should include resource providers of several types (e.g. providers with lower prices and better resources; lower prices and same resources; lower prices and worse resources; and vice versa) and achieve fair allocation of resources.
3. *Collaboration with other providers*: Studies the effect of resource providers collaborating in reducing or increasing their prices, to eliminate competitors at an InterGrid level. In this regard, mechanisms to avoid unexpected behaviour in the economy and the emergence of monopoly or oligopoly would be studied.
4. *Macro-economy*: Each Grid can have its own virtual currency; however, the real value of the currency of Grid A to Grid B is related to the willingness of Grid B in accepting Grid A's currency, and the amount of currency available. The macroeconomic aspects such as different resource prices in each Grid, exchange rate, and inflation also need to be taken into account.

5.4. Infrastructure for Grid economics and estimation of resource requirements

Economic approaches are useful for coping with problems like providing Grid resources to different users with diverging QoS requirements and of how to reward resource suppliers. However, it is not clear whether the Grid economy should use real or virtual currency [87]. Economic models may also require globally trusted entities for several activities such as accounting, usage quota enforcement, and charging. Trust federations [88] would also be required to ensure minimum levels of trust in these entities. Trying to fill the gap of global trust, the International Trust Federation aims at promoting harmonization and synchronization of regional Policy Management Authorities policies.

Although the design of economic institutions for accounting, Grid banking, and charging for resource usage is needed, it may not be possible in practice since it requires interlinking of accounting systems. In addition, if each Grid adopts its own virtual currency, the detailed study of a money exchange system and its impact is important. Furthermore, electronic payment infrastructures for the InterGrid are also difficult, since countries can have different policies regarding the flow of money.

Resource exchange between VOs and Grids is also important. However, it is difficult to agree on standard units for resource usage. For example, it is difficult to evaluate how much storage is equivalent to 30 CPU hours. Moreover, means to compute settlements between Grids through aggregate measurements are needed.

Most of the current economic approaches for Grid computing assume that users know how to estimate their needs for resources, which is a fallacious assumption because users often have trouble estimating their needs [89]. Current approaches assume that the time taken for jobs to execute is known, which is generally not the case in practice. In addition, it is presumed that resource providers know how to estimate the cost for their infrastructure and that capacity planning is not an issue, which is untrue in practice [90]. Thus, these issues should be addressed correctly to effectively apply economic principles in Grids.

5.5. Integration of accounting systems

The integration and connection of accounting systems is necessary in CDI, where multiple CDNs interconnect in order to replicate content from their clients and need means to charge one another either based on the content replicated or the user requests satisfied by each CDN. Similarly, in Grid internetworking the peering between two Grids will require the definition of settlements and in this scenario, the integration of accounting systems is necessary. Grids therefore need to agree upon the measurements for resource usage and provide interoperation of their accounting systems. A work towards this goal is the Extensible and Economics-Inspired Open Grid Computing Platform (EGG) [56]. EGG is a macroeconomics-inspired open Grid platform that promotes the use of resources across Grids through exchange rates, thus allowing each Grid to have their own local currency.

5.6. Management and adaptation of applications

The InterGrid environment requires application models that are able to adapt themselves to the dynamicity of the environment and can report the need of resources to management entities. For example, DVEs can increase or decrease their allocations based on the needs of the applications running on them. However, the flexibility of allocations is dependent on several factors such as cost, time and overhead for changing allocations and underlying peering arrangements. Also, the entire process should be transparent to the applications with little or no degradation in the QoS.

In the InterGrid, resource failures can occur for various reasons: variations in the configuration of the environment, non-availability of required virtual machines or service components, overloaded resource conditions, and faults in computational and network fabric components. In case of failures, alternative resources must be identified and the execution environment migrated. Current check-pointing mechanisms based on the job abstraction may not be enough because the migration of execution environments requires the check-pointing of service components. As the migration of VMs becomes less costly, VM technology becomes appealing to provide the means for migrating execution environments and the recovery from failures. However, these also require new advancements, both in strategies and mechanisms for handling fault tolerance in applications for the InterGrid.

5.7. Peering policies for InterGrid resource allocation and a way forward

Internetworking of Grids and their evolution to the InterGrid is full of challenges. It not only requires fundamental research in business models, methodologies, and mechanisms that enable creation of system for interlinking Grids, but also an open infrastructure that supports standard protocols and interoperability between Grids. As presented in the previous section, it is important to address issues such as price setting of compute resources, InterGrid resource allocation and internetworking of accounting systems. We have been investigating policies for internetworking Grids, particularly for InterGrid resource allocation. We advocate that the peering arrangements and agreements between Grids resemble the manner in which ISPs peer in the Internet. These peering agreements can define, for example: (i) how slot requests are redirected from one Grid to another; (ii) how to balance slot requests preventing an individual Grid from being overloaded while spare capacity exists in peering Grids; and (iii) the common units for resource exchange between Grids.

In the proposed mechanism, Grids can redirect slot requests and balance the load imposed by applications between Grids in a contract network. We consider that Grids establish bilateral contracts between themselves specifying the price for common units of resource—here termed standard slots—allocated from one to another. The mechanism and policies allow the balancing and redirection of requests across Grids. The aim of the mechanism is to prevent an individual Grid from being overloaded while other Grids in the contract network have spare capacity at a price lower than that incurred by the individual Grid. This also resembles the way that power utilities exchange electricity [91], where power plants have varying costs for producing electricity. During peak load periods, the cost for a given power plant to produce electricity crosses a threshold, when it becomes more cost-effective to keep the production at current levels and offload demand to a power plant with lower production cost. Power utilities utilize standard units such as megawatts per hour (MW h) for power exchange.

Moreover, we are also considering a utility-driven model, in which various DVEs can compute their utilities for increasing or decreasing their resource allocations. An IGG allocates resources to DVEs based on the DVEs' utilities, computing the utility for an individual Grid. The need to allocate resources from other Grids, provide resources to other Grids, or claim resources back, is driven by the utility of the individual Grid. Allocation of resources can be constantly moved from one DVE to another in order to maximize the utility at the level of an individual Grid. Some of the challenges that need to be addressed are: (i) obtaining information on resource demand and compute a single value (utility) for each DVE; and (ii) aggregating DVEs' utilities to compute the utility at the level of an individual Grid.

In addition to our effort, the research community has been working to address some issues in these respects, but a massive effort is still needed to realize the internetworking of islands of Grids and develop mechanisms that allow the InterGrid to grow in a sustainable manner. Moreover, the development of applications that can harness the capabilities of such Grids is also a challenge. The design of scalable applications that can utilize the capabilities of networks of Grids is also of utmost importance and a challenging task.

6. RELEVANT WORK

A number of Grid infrastructures have been created around the globe [17–22,25,92]. Recently, there has been interest in promoting interoperability between Grid infrastructures and leveraging best practices in this regard, an example of such work is the effort at the OGF, termed 'Grid Interoperability Now' (GIN). These works focus on various aspects of interoperability, mainly at the middleware level, such as common formats for expressing resources and user groups in Grid Information Systems (GISs), common interfaces for data access, XML-based languages for description of application jobs, common Web Services-based interfaces for job submission and monitoring, among others. While standards and the addressing of above issues are extremely important to guarantee a reasonable level of interoperability between Grids, most of these works do not focus on interlinking of Grids, which involves the peering policies among Grids. Currently, the provisioning, brokering, and scheduling policies are left aside and are not considered in a multi-Grid scenario.

Several architectures for resource management based on leasing of resources have been proposed. Shirako [72], for example, offers an architecture for resource management based on the abstraction of resource leasing. Sites delegate limited power to allocate their resources by registering their resource offerings with brokers. Guest applications can acquire resources from brokers by leasing them for a specified time. In addition, the creation of execution environments has been considered. There are initiatives aiming at the creation and management of execution environments [73]. Although such technology can be leveraged, we need to consider the creation of DVEs spanning multiple Grids and the peering arrangements between Grids. The need for control planes that allow the creation of smaller and autonomous Grids or utility computing infrastructures and the further interlinking of these infrastructures have also been identified [72]. For example, the PlanetLab architecture [68] has been evolving to allow the federation of autonomous PlanetLabs controlled by different organizations [69]. PlanetLab currently provides a global infrastructure and mechanisms that allow the creation of slices; on top of these slices, varying distributed applications can run. The interest in federation, however, will eventually lead to the creation of smaller autonomous PlanetLabs. The use of virtualization technology and federation of autonomous utility infrastructures is also a scenario considered by the GENI [69]. These technologies certainly provide the basis upon which the InterGrid can be built. However, there are many challenges regarding the mechanisms, policies, and application models for the InterGrid.

7. CONCLUSIONS

In this work, we present a case and model for the InterGrid as an evolvable and sustainable system. We start with the analysis of analogous global systems, how they have evolved and what principles can be applied to Grid computing in order to enable the InterGrid vision. We then present an architecture for the InterGrid with the aim of realising it. A discussion on current issues that arise when linking islands of Grids as well as a gap analysis of current Grid technologies was provided to motivate the need for new technologies to enable the InterGrid.

The work contains a discussion on existing projects aiming at creating national and continental Grids. However, many applications currently require amounts of resources that are only achievable by creating Grids of Grids. Existing projects have tried to federate Grids and have provided means to enable VOs to solve several problems. However, the Cyberinfrastructure to cope with these and next-generation challenges will not be realized given that today's Grids follow organizational models and mechanisms that prevent them from internetworking.

Current technologies do not allow the InterGrid vision due to conceptual and technological drawbacks such as the lack of coordination mechanisms. As argued in this work, there is a need for an architecture that allows Grids' structures to evolve from the local to the InterGrid level and enables the easy development of Grid applications for e-Science and e-Business. In addition, many issues related to cultural, social, and political divergences have to be considered or even solved. Like the Internet, the InterGrid will comprise of numerous self-interested stakeholders and the design of its architecture has to consider these aspects [77]. Our contribution in this work is of identifying key problems in realizing a true InterGrid and delineating a research agenda on the topic. As the research agenda is a massive endeavour, we invite the global research community to address some of the issues and work collaboratively in realizing the proposed InterGrid vision.

ACKNOWLEDGEMENTS

The authors are thankful to Krishna Nadiminti, Chee Shin Yeo, and Hussein Gibbins from the University of Melbourne for helping to improve the language and expression in this paper. We also thank Geoffrey Fox (Indiana University), Omer F. Rana (Cardiff University), Jianhong Xia (RMIT University), Fernando Koch (Utretch University), Sushil K. Prasad (Georgia State University), Vanessa Teague (University of Melbourne), Kyong Hoon Kim (University of Melbourne) and Torsten Eymann (Universitt Bayreuth) for sharing their thoughts on the topic. We also thank the anonymous reviewers for their comments and recommendations, which have been crucial to improve the quality of this work.

REFERENCES

1. Baker M, Buyya R, Laforenza D. Grids and Grid technologies for wide-area distributed computing. *Software*: *Practice and Experience* 2002; **32**:1437–1466.
2. Paulson LD. News briefs: Putting a business suit on grid computing. *IEEE Computer* 2005; **38**:24.
3. Abbas A. *Grid Computing*: *A Practical Guide to Technology and Applications*. Charles River Media: Hingham, MA, 2004.
4. Foster I, Kesselman C. Globus: A metacomputing infrastructure toolkit. *International Journal of Supercomputer Applications* 1997; **11**:115–128.
5. Chapin SJ, Katramatos D, Karpovich JF, Grimshaw AS. The Legion resource management system. *Job Scheduling Strategies for Parallel Processing* (*IPPS/SPDP '99/JSSPP '99*), London, U.K., 1999; 162–178.
6. Almond J, Snelling D. UNICORE: Uniform access to supercomputing as an element of electronic commerce. *Future Generation Computer Systems* 1999; **15**:539–548.
7. gLite—Lightweight Middleware for Grid Computing. Available at: http://glite.web.cern.ch/glite, 2005.
8. Berman F, Wolski R. The AppLeS project: A status report. *Eighth NEC Research Symposium*, Berlin, Germany, 1997.
9. Venugopal S, Buyya R, Winton L. A Grid service broker for scheduling e-Science applications on global data grids. *Concurrency and Computation*: *Practice and Experience* 2006; **18**:685–699.
10. Buyya R, Abramson D, Giddy J. Nimrod/G: An architecture for a resource management and scheduling system in a global computational grid. *Fourth International Conference on High Performance Computing in Asia–Pacific Region* (*HPC Asia 2000*), Beijing, China, 2000; 283–289.
11. Frey J, Tannenbaum T, Livny M, Foster I, Tuecke S. Condor-G: A computation management agent for multi-institutional grids. *Cluster Computing* 2002; **5**:237–246.
12. Huedo E, Montero RS, Llorente IM. A framework for adaptive execution in grids. *Software—Practice and Experience* 2004; **34**:631–651.
13. Foster I, Kesselman C, Nick JM, Tuecke S. The physiology of the Grid: An open grid services architecture for distributed systems integration. *Open Grid Service Infrastructure WG*, *Global Grid Forum*, U.S.A., 2002.

14. Freeman PA, Crawford DL, Kim S, Muñoz JL. Cyberinfrastructure for science and engineering: Promises and challenges. *Proceedings of the IEEE* 2005; **93**:682–691.
15. Nandkumar R. International cyberinfrastructure: Activities around the globe. *Cyberinfrastructure Technology Watch Quarterly* 2006; **2**.
16. Emmott S, Rison S. *Towards 2020 Science Report*. Microsoft Research, Cambridge, 2006.
17. Enabling Grids for E-sciencE (EGEE) project. Available at: http://public.eu-egee.org, 2005.
18. Open Science Grid. Available at: http://www.opensciencegrid.org, 2005.
19. Teragrid. Available at: http://www.teragrid.org/, 2006.
20. Catlett C, Beckman P, Skow D, Foster I. Creating and operating national-scale cyberinfrastructure services. *CTWatch Quarterly* 2006; **2**:2–10.
21. K*Grid. Available at: http://www.gridcenter.or.kr/, 2005.
22. Miura K. Overview of Japanese science Grid project NAREGI. *Progress in Informatics* 2006; **3**:67–75.
23. Ram NM, Ramakrishnan S. GARUDA: India's national grid computing initiative. *Cyberinfrastructure Technology Watch Quarterly* 2006; **2**.
24. Hey T, Trefethen AE. The UK e-Science core programme and the grid. *Future Generation Computer Systems* 2002; **18**:1017–1031.
25. Andrade N, Costa L, Germoglio G, Cirne W. Peer-to-peer grid computing with the OurGrid community. *Twenty-third Brazilian Symposium on Computer Networks* (*SBRC 2005*)*—4th Special Tools Session*, 2005.
26. Asia Pacific Grid. Available at: http://www.apgrid.org/, 2005.
27. INWA Project—Australia, China, United Kingdom come together via Grid. *Grid Today—Daily News and Information for the Global Grid Community*, vol. 4, August 2005.
28. Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling scalable virtual organizations. *International Journal of Supercomputer Applications* 2001; **15**(3):200–222.
29. Boghosian B, Coveney P, Dong S, Finn L, Jha S, Karniadakis GE, Karonis NT. Nektar, SPICE and Vortonics: Using federated grids for large scale scientific applications. *Workshop on Challenges of Large Applications in Distributed Environments* (*CLADE*)*—in Conjunction with the 15th International Symposium on High Performance Distributed Computing* (*HPDC-15*), Paris, France, 2006.
30. Chu X, Lonie A, Harris P, Thomas SR, Buyya R. KidneyGrid: A grid platform for integration of distributed kidney models and resources. *Fourth International Workshop on Middleware for Grid Computing* (*MGC 2006*), Melbourne, Australia, 2006; 5.
31. Hunter P, Smith N, Fernandez J, Tawhai M. Integration from proteins to organs: The IUPS physiome project. *Mechanisms of Ageing and Development* 2005; **126**:187–192.
32. Metz C. Interconnecting ISP networks. *IEEE Internet Computing* 2001; **5**:74–80.
33. Kauffman SA. *At Home in the Universe*: *The Search for Laws of Self-organization and Complexity*. Oxford University Press: New York, 1996.
34. Girvan M, Newman MEJ. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America* 2002; **99**:7821–7826.
35. Kumar R, Raghavan P, Rajagopalan S, Sivakumar D. The Web as a graph. *Nineteenth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (*PODS '00*), Dallas, TX, U.S.A., 2000; 1–10.
36. Smarr L. The Grid 2: Blueprint for a New Computing Infrastructure, Foster I, Kesselman C (eds.). Morgan Kaufmann Publishers Inc.: San Francisco, CA, U.S.A., 2003; 3–12.
37. Kurose JF, Ross KW. *Computer Networking*: *A Top-Down Approach Featuring the Internet* (2nd edn). Addison Wesley Professional: Boston, MA, 2002.
38. Norton WB. Internet service providers and peering. *Nineteenth North American Network Operators Group Meeting* (*NANOG 19*), Albuquerque, NM, 2000.
39. Feigenbaum J, Karger DR, Mirrokni VS, Sami R. Subjective-cost policy routing. *First International Workshop on Internet and Network Economics* (*WINE 2005*), Hong Kong, 2005; 174–183.
40. Feigenbaum J, Sami R, Shenker S. Mechanism design for policy routing. *Twenty-third Annual ACM Symposium on Principles of Distributed Computing* (*PODC '04*), New York, NY, U.S.A., 2004; 11–20.
41. Baake P, Wichmann T. On the economics of Internet peering. *NETNOMICS* 1999; **1**:89–105.
42. Huston G. Interconnection, peering and settlements—Part I. *The Internet Protocol Journal* 1999; **2**:2–16.
43. Berners-Lee T, Fischetti M. *Weaving the Web*: *The Past*, *Present and Future of the World Wide Web by its Inventor*. Orion Business: London, 1999.
44. Timmers P. Business models for electronic markets. *Journal on Electronic Markets* 1998; **8**:3–8.
45. Dilley J, Maggs B, Parikh J, Prokop H, Sitaraman R, Weihl B. Globally distributed content delivery. *IEEE Internet Computing* 2002; **6**:50–58.
46. Mirror Image Internet. About the company. Available at: http://www.mirror-image.com/company, 2006.
47. Amini L, Shaikh A, Schulzrinne H. Effective peering for multi-provider content delivery services. *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, Hong Kong, 2004; 850–861.

DOI: 10.1002/cpe

48. Rzewski P, Day MS, Gilletti D. Content Internetworking (CDI) scenarios. *RFC 3570*, IETF, 2003.
49. Day MS, Cain B, Tomlinson G, Rzewski P. A model for Content Internetworking (CDI). *RFC 3466*, IETF, 2003.
50. Economides N. The economics of networks. *International Journal of Industrial Organization* 1996; **16**:673–699.
51. Blake S, Black DS, Carlson MA, Davies E, Wang Z, Weiss W. An architecture for differentiated services. *RFC 2475*, IETF, 1998.
52. Nichols K, Carpenter B. Definition of differentiated services per domain behaviors and rules for their specification. *RFC 3086*, IETF, 2001.
53. Black DL, Brim SW, Carpenter BE, Faucheur FL. Per hop behavior identification codes. *RFC 3140*, IETF, 2001.
54. Buyya R. *Economic-based Distributed Resource Management and Scheduling for Grid Computing*. Monash University: Melbourne, Australia, 2002.
55. Lai K, Rasmusson L, Adar E, Sorkin S, Zhang L, Huberman BA. Tycoon: An implementation of a distributed market-based resource allocation system. *Technical Report*, HP Labs, Palo Alto, CA, U.S.A., 2004.
56. Brunelle J, Hurst P, Huth J, Kang L, Ng C, Parkes D, Seltzer M, Shank J, Youssef S. EGG: An extensible and economics-inspired open grid computing platform. *Third International Workshop on Grid Economics and Business Models* (*GECON 2006*), Singapore, 2006; 140–150.
57. Grace P, Coulson G, Blair G, Mathy L, Yeung WK, Cai W, Duce D, Cooper C. GRIDKIT: Pluggable overlay networks for Grid computing. *Distributed Objects and Applications* (*DOA '04*), Cyprus, 2004.
58. Ranjan R, Buyya R, Harwood A. A case for cooperative and incentive-based coupling of distributed clusters. *Seventh IEEE International Conference on Cluster Computing*, Boston, MA, U.S.A., 2005.
59. Anastasiou M, Bifulco A, Findeisen P, Hannus M, Karvonen I, Löh H, Plüss A, Ollus M, Weidemann M. VO Guidelines Draft v.19. *Deliverable D54.1*, Virtual Organizations Cluster, 2004.
60. Kwok Y-K, Song S, Hwang K. Selfish Grid computing: Game-theoretic modeling and NAS performance results. *International Symposium on Cluster Computing and the Grid* (*CCGrid '05*), Cardiff, U.K., 2005.
61. Dumitrescu C, Wilde M, Foster I. A model for usage policy-based resource allocation in Grids. *Sixth IEEE International Workshop on Policies for Distributed Systems and Networks*, 2005; 191–200.
62. Alfieri R, Cecchini R, Ciaschini V, dell'Agnello L, Frohner Á, Lörentey K, Spataro F. From gridmap-file to VOMS: Managing authorization in a Grid environment. *Future Generation Computer Systems* 2005; **21**:549–558.
63. Nasser B, Benzekri A, Laborde R, Grasset F, Barrère F. Access control model for Grid virtual organizations. *Seventh International Conference on Enterprise Information Systems* (*ICEIS 2005*), Miami, FL, U.S.A., 2005; 152–158.
64. Carpenter BE, Janson PA. Abstract interdomain security assertions: A basis for extra-Grid virtual organizations. *IBM Systems Journal* 2004; **43**:689–701.
65. Hayek FAv. *The Fatal Conceit*: *The Errors of Socialism*, vol. 1. Routledge: London, 1988.
66. De Roure D. On self-organization and the semantic grid. *IEEE Intelligent Systems* 2003; **18**:77–79.
67. Ramakrishnan L, Grit L, Iamnitchi A, Irwin D, Yumerefendi A, Chase J. Toward a doctrine of containment: Grid hosting with adaptive resource control. *ACM/IEEE Supercomputing 2006 Conference* (*SC '06*), Tampa, U.S.A., 2006; 20–20.
68. Peterson L, Muir S, Roscoe T, Klingaman A. PlanetLab architecture: An overview. *PlanetLab Consortium*, Princeton, NJ, U.S.A., May 2006.
69. Peterson L, Wroclawski J. Overview of the GENI architecture. *GENI Design Document GDD-06-11*, GENI: Global Environment for Network Innovations, January 2007.
70. Barham P, Dragovic B, Fraser K, Hand S, Harris T, Ho A, Neugebauer R, Pratt I, Warfield A. Xen and the art of virtualization. *Nineteenth ACM Symposium on Operating Systems Principles*, Bolton Landing, NY, U.S.A., 2003; 164–177.
71. Keahey K, Foster I, Freeman T, Zhang X. Virtual workspaces: Achieving quality of service and quality of life in the Grid. *Scientific Programming* 2005; **13**(4):265–275.
72. Irwin D, Chase J, Grit L, Yumerefendi A, Becker D, Yocum KG. Sharing networked resources with brokered leases. *USENIX Technical Conference*, Boston, MA, 2006.
73. Ruth P, Jiang X, Xu D, Goasguen S. Virtual distributed environments in a shared infrastructure. *IEEE Computer* 2005; **38**:63–69.
74. Adabala S, Chadha V, Chawla P, Figueiredo R, Fortes J, Krsul I, Matsunaga A, Tsugawa M, Zhang J, Zhao M, Zhu L, Zhu X. From virtualized resources to virtual computing Grids: The InVIGO system. *Future Generation Computer Systems* 2005; **21**:896–909.
75. Shoykhet A, Lange J, Dinda P. Virtuoso: A system for virtual machine marketplaces. *Technical Report*, Electrical Engineering and Computer Science Department, Northwestern University, Evanston/Chicago, IL, July 2004.
76. Sundararaj AI, Gupta A, Dinda PA. Increasing application performance in virtual environments through run-time inference and adaptation. *Fourteenth IEEE International Symposium on High Performance Distributed Computing* (*HPDC 2005*), Research Triangle Park, NC, 2005; 47–58.
77. Clark DD, Wroclawski J, Sollins KR, Braden R. Tussle in Cyberspace: Defining tomorrow's Internet. *IEEE/ACM Transactions on Networking* 2005; **13**:462–475.

DOI: 10.1002/cpe

78. Sundararaj AI. Automatic, run-time and dynamic adaptation of distributed applications executing in virtual environments. *PhD Dissertation*, *Technical Report NWU-EECS-06-18*, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston/Chicago, IL, 2006.
79. Singh A, Haahr M. Creating an adaptive network of hubs using Schelling's model. *Communications of the ACM* 2006; **49**:69–73.
80. Heylighen F. Knowledge management, organizational intelligence and learning, and complexity. *The Science of Self-organization and Adaptivity*. Eolss Publishers Co. Ltd.: Oxford, U.K., 2003.
81. Pattnaik P, Ekanadham K, Jann J. Autonomic computing and Grid. *Grid Computing*: *Making the Global Infrastructure a Reality*, Berman F, Fox G, Hey T (eds.). Wiley: New York, 2003; 351–361.
82. Weiss MB, Shin SJ. Internet interconnection economic model and its analysis: Peering and settlement. *NETNOMICS* 2004; **6**:43–57.
83. Ranjan R, Buyya R, Harwood A. A model for cooperative federation of distributed clusters. *Fourteenth IEEE International Symposium on High Performance Distributed Computing* (*HPDC 2005*), Research Triangle Park, NC, U.S.A., 2005.
84. Balakrishnan H, Shenker S, Walfish M. Peering peer-to-peer providers. *Fourth International Workshop Peer-to-Peer Systems* (*IPTPS 2005*), Ithaca, NY, U.S.A., 2005; 104–114.
85. Badasyan N, Chakrabarti S. Private peering, transit and traffic diversion. *NETNOMICS* 2005; **7**:115–124.
86. Buyya R, Abramson D, Venugopal S. The Grid economy. *Special Issue on Grid Computing*, *Proceedings of the IEEE* 2005; **93**:698–714.
87. Sandholm T, Gardfjäll P, Elmroth E, Johnsson L, Mulmo O. An OGSA-based accounting system for allocation enforcement across HPC centers. *Second International Conference on Service Oriented Computing* (*ICSOC '04*), New York, NY, U.S.A., 2004; 279–288.
88. International Grid Trust Federation. Available at: http://www.gridpma.org/, 2005.
89. Bhargava HK, Sundaresan S. Contingent bids in auctions: Availability, commitment and pricing of computing as utility. *Thirty-seventh Annual Hawaii International Conference on System Sciences* (*HICSS '04*), Washington, DC, U.S.A., 2004; 80217.3.
90. Krishnan R, Hosanagar K. Challenges in designing grid marketplaces. *Third International Workshop on Grid Economics and Business Models* (*GECON 2006*), Singapore, 2006; 61–69.
91. Larsson S. Energy transfers between Scandinavian countries. *IEE Colloquium on System Interconnection and Energy Exchange Across National Boundaries*, London, U.K., 1992; 5/1-513.
92. Australian Partnership for Advanced Computing—GRID. Available at: http://www.apac.edu.au/programs/GRID/index.html, 2005.