(Hoffa, Mehta et al. 2008; House, Marshall et al. 2008; Kacsuk, Kertesz et al. 2008; Lana and Westphall 2008; Litke, Konstanteli et al. 2008; Maoz, Barak et al. 2008; Montes, Sanchez et al. 2008; Opitz, KÃ¶nig et al. 2008; Palankar, Lamnitchi et al. 2008; Pandey, Chao et al. 2008; Pugliese, Talia et al. 2008; Rahman and Buyya 2008; Reed 2008; Richard, Joshi et al. 2008; Satayapiwat, Egawa et al. 2008; Song, Li et al. 2008; Stratan, Iosup et al. 2008; VÃ¡zquez, Huedo et al. 2008; Wu, Deng et al. 2008; Wu, Zhu et al. 2008; Xie, Huang et al. 2008; Youseff, Butrico et al. 2008; Yu, Buyya et al. 2008; Anandasivam and Neumann 2009; Erdogmus 2009; Vanderster, Dimopoulos et al. 2009; Vaquero, Rodero-Merino et al. 2009; Vykoukal, Wolf et al. 2009) (Alkadi and Alkadi 2006; Azzedin, Maheswaran et al. 2006; Darlington, Cohen et al. 2006; Gentzsch 2006; Kawato, Machida et al. 2006; Liu, Wang et al. 2006; Llorente, Montero et al. 2006; Sandholm, Lai et al. 2006; Shaheen, Malik et al. 2006; Smith, Engel et al. 2006; Tang, Li et al. 2006; Truong, Samborski et al. 2006; Yeo and Buyy 2006; Yuan, Zeng et al. 2006; Zhang, Liu et al. 2006; Zhang, Lin et al. 2006; Caracas and Altmann 2007; Denne 2007; Franke, Theilmann et al. 2007; Garbacki and Naik 2007; Ivanov 2007; Laccetti and Schmid 2007; Li and Li 2007; Li and Li 2007; Losup, Epema et al. 2007; Narmadha and Selvi 2007; Ogawa, Itoh et al. 2007; Olugbara, Adigun et al. 2007; Ouyang, Sahai et al. 2007; Quétier, Neri et al. 2007; Samimi, McKinley et al. 2007; Song, Sun et al. 2007; Soni and Sharma 2007; Stockinger 2007; Vazquez, Huedo et al. 2007; Yu, Zheng et al. 2007; Allenotor and Thulasiram 2008; Arustei, Craus et al. 2008; Aymerich, Fenu et al. 2008; Beck, Schwind et al. 2008; Buyya, Yeo et al. 2008; Cendron and Westphall 2008; Cohen, Darlington et al. 2008; de Assuncao, Buyya et al. 2008; Ditarso, Figueiredo et al. 2008; Dube and Parizeau 2008; Dun, Taura et al. 2008; Foster, Zhao et al. 2008; GjermundoÃ¸r, Dikaiakos et al. 2008)(Anderson 1996; Foster and Kesselman 1998; Foster and Kesselman 1999; Leinberger and Kumar 1999; Buyya, Abramson et al. 2000; Foster and Kesselman 2000; Ferguson and Towns 2001; Baker, Buyya et al. 2002; Chen 2002; Chase, Irwin et al. 2003; Machiraju, Sahai et al. 2003; Trezentos and Oliveira 2003; Venugopal, Buyya et al. 2003; Yongwei, Guangwen et al. 2003; Bhargava and Sundaresan 2004; Buyya and Venugopal 2004; Ernemann, Hamscher et al. 2004; Gagliardi 2004; Maheswaran, Maniymaran et al. 2004; Maniymaran and Maheswaran 2004; Menasce and Casalicchio 2004; Sahai, Singhal et al. 2004; Weiss 2004; Zanikolas and Sakellariou 2004; Zhang, Li et al. 2004; Buyya, Abramson et al. 2005; Byun, Jang et al. 2005; de Assuncao, Nadiminti et al. 2005; Gagliardi, Jones et al. 2005; Ghosh, Roy et al. 2005; Jiang, Meng et al. 2005; Komisarczuk and Bubendorfer 2005; Li, Yuanzhuo et al. 2005; Liang-Jie and Zongwei 2005; Medernach 2005; Obozinski 2005; Peng, Ng et al. 2005; van Moorsel 2005; Yu and Buyya 2005; Yu, Buyya et al. 2005)

Alkadi, I. and G. Alkadi (2006). "Grid computing: The past, now, and future." Human Systems Management **25**(3): 161-166.

With the speed of changing technology in our time, staying up to date is very hard, and sometimes impossible for companies. Companies have to constantly update computers, servers, databases, and applications, among other things. Even though computers need to be updated to handle sophisticated software and complex problems, there is still vast amounts of power going to waste by computers sitting idle. A new technology has been formed out of this idea to take advantage of the amount of power being unused to allow companies to make use of it, saving vast amounts of much needed company funds. Grid Computing is a relatively new approach to sharing and accessing resources such as, computers, storage devices, and peripherals, which

goes far beyond distributed computing. Accomplishing the goals of Grid Computing, however, present many challenges as well as opportunities. Some of these challenges include (a) security, (b) programming, (c) proper allocation of resources, and (d) inter processsystem communication, to name a few. Currently, most successful deployment of Grid Computing exists in the areas of Physics and Scientific research; however, very little has been done in business and organizations. In this paper, we discuss issues and exploit possibilities for grid deployment in organizations, with a special emphasis on how such deployment may help to improve efficiency as well as determining an acceptable on return on the required level of investment in this new technology. © 2006 - IOS Press and the authors. All rights reserved.

Allenotor, D. and R. K. Thulasiram (2008). Grid resources pricing: A novel financial option based quality of service-profit quasi-static equilibrium model. Proceedings - IEEE/ACM International Workshop on Grid Computing, Tsukuba.

Use of grid resources has been free so far and a trend is developing to charge the users. The challenges that characterize a grid resource pricing model include the dynamic ability of the model to provide a high satisfaction guarantee measured as Quality of Service (QoS) - from users perspectives, profitability constraints - from the grid operator perspectives, and the ability to orchestrate grid resources for their availability on-demand. In this study, we design, develop, and simulate a grid resources pricing model that balances these constraints. We employ financial option theory and treat the grid resources as real assets to capture the realistic value of the grid compute commodities. We then price the grid resources by solving the finance model. We discuss the results on pricing of compute cycles based on the actual data of grid usage pattern obtained from the WestGrid and the SHARCNET. We extend and generalize our study to any computational grid. Â© 2008 IEEE.

Anandasivam, A. and D. Neumann (2009). Managing Revenue in Grids. System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on.

The distributed usage of computing resources over a large-scale network allows users to receive and offer resources on demand. The on demand paradigm leads to dynamic and unpredictable usage of resources, since every user in the network will try to maximize his utility by selfish behavior. The customer's behavior can be actuated by pricing policies to lower demand at peak time. Revenue Management as a relatively new economic paradigm provides various tools to optimally allocate capacity and increase revenue. We provide a framework how the matured concepts of Revenue Management can be deployed to Grid Computing. We analyze whether the Grid Computing domain has notable differences from the airline industry or other common areas for Revenue Management like restaurant, hotel or car rental industries. Hence, we outline tools and methods known from Revenue Management and how they can be applied to Grid Computing.

Anderson, C. (1996). "Western Kentucky makes computing a basic utility." Communications News **33**(5): 26-27.

Arustei, S., M. Craus, et al. (2008). Towards a generic framework for deploying applications as grid services. Proceedings of the 2nd International Workshop on Architectures, Concepts and Technologies for Service Oriented Computing, ACT4SOC 2008 - In Conjunction with ICSOFT 2008, Porto.

Exploiting the power of the grid very often involves transforming existing or new applications into grid services. In this paper we present a generic framework based on a service oriented architecture developed in order to simplify the task of deploying

applications as grid services. Our work consists of a configurable grid service that provides application developers with a high level programming model, hiding the complexity of dealing with web services and grid technologies. The architectural design of the framework allows custom functionality to be plugged into an adaptive grid service in a simple manner, thus attracting more non-expert users to the grid. A prototype implementation of the framework has been built and a case study has been developed to illustrate the concept.

Aymerich, F. M., G. Fenu, et al. (2008). An Approach to a Cloud Computing Network. 1st International Conference on the Applications of Digital Information and Web Technologies, Ostrava, CZECH REPUBLIC, Ieee.

"Cloud Computing" is becoming increasingly relevant, as it will enable companies involved in spreading this technology to open the doors to Web 3.0. In this work the basic features of cloud computing are presented and compared with those of the original technology: Grid Computing. The new categories of services introduced will slowly replace many types of computational resources currently used. In this perspective, grid computing, the basic element for the large scale supply of cloud services, will play a fundamental role in defining how those services will be provided. The paper describes the concept of computational resources outsourcing, referred to computational grids and a real application. This work utilises the results by the Cybersar Project managed by the COSMOLAB Consortium (Italy).

Azzedin, F., M. Maheswaran, et al. (2006). "Trust brokering and Its use for resource matchmaking in public-resource grids." Journal of Grid Computing **4**(3): 247-263.

This paper presents a trust brokering system that operates in a peer-to-peer manner. The network of trust brokers operate by providing peer reviews in the form of recommendations regarding potential resource targets. One of the distinguishing features of our work is that it separately models the accuracy and honesty concepts. By separately modeling these concepts, our model is able to significantly improve the performance. We apply the trust brokering system to a resource manager to illustrate its utility in a public-resource Grid environment. The simulations performed to evaluate the trust-aware resource matchmaking strategies indicate that high levels of 'robustness' can be attained by considering trust while matchmaking and allocating resources. Â© Springer Science+Business Media B.V. 2006.

Baker, M., R. Buyya, et al. (2002). "Grids and Grid technologies for wide-area distributed computing." Software-Practice & Experience **32**(15): 1437-1466.

The last decade has seen a substantial increase in commodity computer and network performance, mainly as a result of faster hardware and more sophisticated software. Nevertheless, there are still problems, in the fields of science, engineering, and business, which cannot be effectively dealt with using the current generation of supercomputers. In fact, due to their size and complexity, these problems are often very numerically and/or data intensive and consequently require a variety of heterogeneous resources that are not available on a single machine. A number of teams have conducted experimental studies on the cooperative use of geographically distributed resources unified to act as a single powerful computer. This new approach is known by several names, such as metacomputing, scalable computing, global computing, Internet computing, and more recently peer-to-peer or Grid computing. The early efforts in Grid computing started as a project to link supercomputing sites, but have now grown far beyond their original intent. In fact, many applications can benefit from the Grid infrastructure, including collaborative engineering, data exploration, high-throughput computing, and of course distributed supercomputing.

Moreover, due to the rapid growth of the Internet and Web, there has been a rising interest in Web-based distributed computing, and many projects have been started and aim to exploit the Web as an infrastructure for running coarse-grained distributed and parallel applications. In this context, the Web has the capability to be a platform for parallel and collaborative work as well as a key technology to create a pervasive and ubiquitous Grid-based infrastructure. This paper aims to present the state-of-the-art of Grid computing and attempts to survey the major international efforts in developing this emerging technology. Copyright (C) 2002 John Wiley Sons, Ltd.

Beck, R., M. Schwind, et al. (2008). "Grid economics in departmentalized enterprises." Journal of Grid Computing **6**(3): 277-290.

The application of Grid technology is finally spreading from engineering and natural science-related industrial sectors to other industries with a high demand for computing applications. However, the diffusion of Grid technology within these sectors is often hindered by a lack of the incentive to share the computational resources across departments or branches even within the same enterprise. A promising way of overcoming these barriers is the introduction of a pricing mechanism for the use of Grid-based resources. This work introduces such a pricing approach to Grid computing and provides three simulation scenarios to illustrate the effectiveness of such an economized Grid solution. The simulation results indicate that the pooling of information technology resources can produce a reduction of 33% in cost compared to individual and dedicated servers. However, with a price-based allocation of computing resources, a further 10% of cost reduction can be achieved by introducing an auction mechanism. Therefore we claim that there is huge cost reduction potential in departmentalized enterprises beyond the savings that can be achieved by a utility-based allocation of computing resources, if economically measured allocation methods are combined with advanced refining and learning methods in the allocation process. Â© Springer Science+Business Media B.V. 2008.

Bhargava, H. K. and S. Sundaresan (2004). "Computing as utility: Managing availability, commitment, and pricing through contingent bid auctions." Journal of Management Information Systems **21**(2): 201-227.

Enabled by advances in grid and network computing architectures for the delivery of on-demand computing services, the vision of an e-services economy in which computing will be as ubiquitous as a utility is becoming a possibility in business computing. Major firms in the computing industry such as IBM, Hewlett-Packard, and Sun Microsystems are focusing on agility and flexibility of computing resources and gearing up for their own versions of on-demand computing and information technology (IT) outsourcing solutions. The successful introduction of these new computing models requires the development of appropriate pricing mechanisms that are consistent with the enabling technologies. Our paper introduces the notion of contingent auctions to address this lacuna. In contingent auctions, users bid for computing resources in an auction, but are relieved from the contract (paying a penalty) if demand is not realized. We study different mechanisms-ranging from an advance commitment (capacity reservation) to no commitment (pay-as-you-go)-under demand uncertainty. We consider markets in which the demand for computing is uncertain and, moreover, users' value of computing and demand realization may be related. We show how the different levels of commitment affect prices, revenues, and resource utilization under different market conditions. Our results reiterate the need to address the availability-commitment dichotomy in the design of business models for on-demand computing and IT outsourcing.

Buyya, R., D. Abramson, et al. (2000). An economy driven resource management architecture for global computational power grids. International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2000), Las Vegas, Nv, C S R E a Press.

The growing computational power requirements of grand challenge applications has promoted the need for linking highperformance computational resources distributed across multiple organisations. This is fueled by the availability of the Internet as a ubiquitous commodity communication media, low cost high-performance machines such as clusters across multiple organisations, and the rise of scientific problems of multi-organisational interest. The availability of expensive, special class of scientific instruments or devices and data sources in few organisations has increased the interest in offering a remote access to these resources. The recent popularity of coupling (local and remote) computational resources, special class of scientific instruments, and data sources across the Internet for solving problems has led to the emergence of a new platform called "Computational Grid". This paper identifies the issues in resource management and scheduling driven by computational economy in the emerging grid computing context. They also apply to clusters of clusters environment (known as federated clusters or hyperclusters) formed by coupling multiple (geographically distributed) clusters located in the same or different organisations. We discuss our current work on the Nimrod/G resource broker, whose scheduling mechanism is driven by a user supplied application deadline and a resource access budget. However, current Grid access frameworks do not provide the dynamic resource trading services that are required to facilitate flexible application scheduling. In order to overcome this limitation, we have proposed an infrastructure called GRid Architecture for Computational Economy (GRACE). In this paper we present the motivations for grid computing, resource management architecture, Nimrod/G resource broker, computational economy, and GRACE infrastructure and its APIs along with future work.

Buyya, R., D. Abramson, et al. (2005). "The grid economy." Proceedings of the IEEE **93**(3): 698-714.

This paper identifies challenges in managing resources in a Grid computing environment and proposes computational economy as a metaphor for effective management of resources and application scheduling. It identifies distributed resource management challenges and requirements of economy-based Grid systems, and discusses various representative economy-based systems, both historical and emerging, for cooperative and competitive trading of resources such as CPU cycles, storage, and network bandwidth. It presents an extensible, service-oriented Grid architecture driven by Grid economy and an approach for its realization by leveraging various existing Grid technologies. It also presents commodity and auction models for resource allocation. The use of commodity economy model for resource management and application scheduling in both computational and data grids is also presented. Â© 2005 IEEE.

Buyya, R. and S. Venugopal (2004). The gridbus toolkit for service oriented grid and utility computing: An overview and status report. GECON 2004 - 2004 1st IEEE International Workshop on Grid Economics and Business Models, Seoul.

Grids aim at exploiting synergies that result from cooperation of autonomous distributed entities. The synergies that result from grid cooperation include the sharing, exchange, selection, and aggregation of geographically distributed resources such as computers, data bases, software, and scientific instruments for solving large-scale problems in science, engineering, and commerce. For this cooperation to be sustainable, participants need to have economic incentive. Therefore, "incentive"

mechanisms should be considered as one of key design parameters of Grid architectures. In this article, we present an overview and status of an open source Grid toolkit, called Gridbus, whose architecture is fundamentally driven by the requirements of Grid economy. Gridbus technologies provide services for both computational and data grids that power the emerging eScience and eBusiness applications.

Buyya, R., C. S. Yeo, et al. (2008). Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities. 10th IEEE International Conference on High Performance Computing and Communications, Dalian, PEOPLES R CHINA, Ieee Computer Soc.

This keynote paper: presents a 21(st) century vision of computing; identifies various computing paradigms promising to deliver the vision of computing utilities defines Cloud computing and provides the architecture for creating market-oriented Clouds by leveraging technologies such as VMs; provides thoughts on market-based resource management strategies that encompass both customer-driven service management and computational risk management to sustain SLA-oriented resource allocation; presents some representative Cloud platforms especially those developed in industries along with our current work towards realising market-oriented resource allocation of Clouds by leveraging the 3(rd) generation Aneka enterprise Grid technology; reveals our early thoughts on interconnecting Clouds for dynamically creating an atmospheric computing environment along with pointers to future community research; and concludes with the need for convergence of competing IT paradigms for delivering our 21(st) century vision.

Byun, E. K., J. W. Jang, et al. (2005). A dynamic grid services deployment mechanism for on-demand resource provisioning. 2005 IEEE International Symposium on Cluster Computing and the Grid, CCGrid 2005, Cardiff, Wales.

Recently Grid computing has started to leverage Web services technology by proposing OGSI-standard. OGSI standard defines the Grid service which presents unified interfaces to every participant of Grid. In current Glubus Toolkit3(GT3),which is an implementation of OGSI, Grid service factories should be deployed manually into resources to provide grid services. However, it is necessary to dynamically allocate proper amount of resource, since the demand for resource of service provider changes over time. In this paper, we propose a architecture to enable on-demand resource provisioning. We develop Universal Factory Service (UFS) that provides a dynamic Grid service deployment mechanism and a resource broker called Door service. Through the experiments, we show that Grid services can adoptively exploit resources according to the request rates. Â© 2005 IEEE.

Caracas, A. and J. Altmann (2007). A pricing information service for grid computing, Newport Beach, CA, United states, IEEE Computer Society.

This paper addresses two shortcomings that exist in the area of pricing Grid services in an economic Grid environment. The first shortcoming is that there are no standards for pricing schemes, caused by a large difference in the units that are traded (e.g. CPU cycles or virtual clusters) in Grid computing. The second shortcoming is the lack of a model for managing the pricing of informational elements (e.g. software applications) and computational elements (e.g. virtual machines, which comprise resources such as CPU, memory, disk space, network bandwidth). This paper presents a pricing service for Grid computing services, which resolves the shortcomings by introducing a general pricing scheme for informational and computational elements. We describe the functional requirements, architecture, and the interfaces of the pricing service. The pricing service allows expressing the proposed general pricing scheme as an XML

document, which can be linked to service level agreements. Contrary to other proposals on pricing, the pricing service is separated from the functionality of metering, accounting, and payment. To validate the concept of a pricing information service, we portray two Utility Computing scenarios. Copyright 2007 ACM.

Cendron, M. M. and C. B. Westphall (2008). A price-based task scheduling for grid computing. Proceedings - 7th International Conference on Networking, ICN 2008, Cancun.

In this paper, we present some principles of grid economy, why it's useful to scheduling task and what is involved in a Grid computing. And we present some results obtained with a framework that use the two model of economy to form a price of a service based, in a macroeconomic way is attributed a initial price for a resource based in your computation characteristics, this is a commodities market, and the microeconomic way, with the node varying your local price based on the demand/supply economic model. Â© 2008 IEEE.

Chase, J. S., D. E. Irwin, et al. (2003). Dynamic virtual clusters in a grid site manager, Los Alamitos, CA, USA, IEEE Comput. Soc.

This paper presents new mechanisms for dynamic resource management in a cluster manager called Cluster-on-Demand (COD). COD allocates servers from a common pool to multiple virtual clusters (vclusters), with independently configured software environments, name spaces, user access controls, and network storage volumes. We present experiments using the popular Sun GridEngine batch scheduler to demonstrate that dynamic virtual clusters are an enabling abstraction for advanced resource management in computing utilities and grids. In particular, they support dynamic, policy-based cluster sharing between local users and hosted Grid services, resource reservation and adaptive provisioning, scavenging of the idle resources, and dynamic instantiation of Grid services. These goals are achieved in a direct and general way through a new set of fundamental cluster management functions, with minimal impact on the Grid middleware itself

Chen, T. M. (2002). "Internet computing as a utility - An invited commentary." Ieee Communications Magazine **40**(5): 10-10.

Cohen, J., J. Darlington, et al. (2008). "Payment and negotiation for the next generation Grid and Web." Concurrency and Computation: Practice and Experience **20**(3): 239-51.

We present a proposal for a next-generation Internet based on chargeable Web Services and Utility Computing realized by a series of open but interacting markets. We demonstrate through the U.K. e-Science project 'A Market for Computational Services' the development of some of the fundamental building blocks for such a Grid computational marketplace. This paper describes the motivation behind this restructuring of the Internet and Web-based activities as a series of markets and how Grid Computing technologies can contribute towards this goal. The paper details the work undertaken at the London e-Science Centre to build a framework to create and support negotiable and chargeable Web Services. Copyright 2007 John Wiley Sons, Ltd.

Darlington, J., J. Cohen, et al. (2006). An architecture for a next-generation internet based on web services and utility computing. Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WET ICE, Manchester.

We present an architecture for a next-generation Internet based on Web Services and Utility Computing. This architecture would be realised as a series of open but interacting markets. We present a scenario whereby these markets could develop, driven by economic forces and market opportunities, exploiting technologies

developed in eScience and Grid computing. We argue that, if properly realised, such an architecture would provide a safer and more cost-effective way of supporting the Internet than the current infrastructure and would provide many opportunities for radical developments, affecting and unifying many areas of academia, business and society. Â© 2006 IEEE.

de Assuncao, M. D., R. Buyya, et al. (2008). "InterGrid: a case for internetworking islands of Grids." Concurrency and Computation-Practice & Experience **20**(8): 997-1024.

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 (C) 2007 John Wiley & Sons, Ltd.

de Assuncao, M. D., K. Nadiminti, et al. (2005). An integration of global and enterprise grid computing: Gridbus Broker and Xgrid perspective. International Symposium on Trustworthy Global Computing, Edinburgh, SCOTLAND, Springer-Verlag Berlin.

In this work, we present two perspectives of Grid computing by using two different Grid middleware as examples: an Enterprise Grid using Xgrid and a Global Grid with Gridbus. We also present the integration of Enterprise and Global Grids by proposing the integration of Gridbus Broker with diverse enterprise Grid middleware including Xgrid, PBS, Condor and SGE. The sample performance results demonstrate the usefulness of this integration effort.

Denne, M. (2007). Pricing utility computing services. 1st International Workshop on Services Engineering, Melbourne, AUSTRALIA, Idea Group Publishing.

Utility computing is the emerging term for the delivery of information technology in a "pay-as-you-go" model. It has attracted considerable attention as a means of delivering lower total cost of ownership (TCO) and more predictable service levels for in-house IT At its heart, utility transformation allows IT consumers to switch from capital-based procurement of IT assets to operational cost procurement of IT services. Unsurprisingly, the change is closely linked with the adoption of service oriented architectures (SOA) and service oriented infrastructures (SOI). In fact, utility transformation often provides the overarching business and financial framework for driving a move to SOA. It defines the IT chargeback environment and the resulting compelling business context. Key to the success of utility transformation is the implementation of appropriate service pricing models. A variety of innovative pricing models can be used to improve service predictability, to create incentives for certain behaviors, and to manage the flow of notional revenue to the IT organization (ITO).

They are invaluable to the SOA business case. This article examines several such models.

Ditarso, P., F. Figueiredo, et al. (2008). On the planning of a hybrid IT infrastructure. NOMS 2008 - IEEE/IFIP Network Operations and Management Symposium: Pervasive Management for Ubiquitous Networks and Services, Salvador - Bahia.

With the emergence of utility computing and the continuous search for reducing the cost of running Information Technology (IT) infrastructures, we will soon experience an important change on the way these infrastructures are assembled, configured and managed. In this paper we consider the problem of managing a hybrid high-performance computing infrastructure whose processing elements comprise in-house dedicated machines, a utility computing service provider, and idle machines from a best-effort peer-to-peer grid. This infrastructure supports the execution of both best-effort and real-time applications. Real-time applications use primarily computing power from the inhouse machines and any processing power that can be attained from the best-effort grid. Extra capacity required to meet deadlines is purchased from the utility computing service provider. This extra capacity is reserved for future use through short term contracts which are negotiated with no human intervention. We take a business-driven approach for the management of this hybrid infrastructure and propose heuristics that can be used by a contract planner agent to reduce the cost of running the applications at the same time that guarantees that deadlines are met. In particular, we show that constructing an estimation for the behavior of the grid is essential for making contracts that lead to high efficiency in the use of the hybrid infrastructure. ©2008 IEEE.

Dube, N. and M. Parizeau (2008). Utility computing and market-based scheduling: Shortcomings for Grid resources sharing and the next steps. Proceedings - International Symposium on High Performance Computing Systems and Applications, Quebec, QC.

This paper lays out fundamental social and economic principles that must be addressed in order to enable true High-Performance Computing resources sharing at the scale of a national Grid. Focusing on the literature around the concept of a Grid Economy, it starts by a rather exhaustive survey of related work in utility computing, marketbased scheduling and other economic approaches in resource allocation. Building on forces and weaknesses of previous work, high-level principles to be followed to establish a sustainable and pervasive Grid sharing framework are detailed. Â© 2008 IEEE.

Dun, N., K. Taura, et al. (2008). GMount: Build your grid file system on the fly. Proceedings - IEEE/ACM International Workshop on Grid Computing, Tsukuba.

By GMount, non-privilege users can easily and quickly build ad-hoc distributed file systems on any machines reachable via SSH. In the wide-area Grid environments, it can scale to hundreds of nodes and works with NAT or firewall. Given the network topology, the metadata operations of file system are locality-aware. GMount can be effortlessly deployed in multiple clusters without superuser privilege. In this paper, we present the design and implementation of GMount, and shows its viability in a large scale Grid platform with over 300 nodes spread across 11 clusters. Â© 2008 IEEE.

Erdogmus, H. (2009). "Cloud Computing: Does Nirvana Hide behind the Nebula?" Software, IEEE **26**(2): 4-6.

At the core of cloud computing is a simple concept: software as a service, or SaaS. Whether the underlying software is an application, application component, platform, framework, environment, or some other soft infrastructure for composing applications

to be delivered as a service on the Web, it's all software in the end. But the simplicity ends there. Just a step away from that core, a complex concoction of paradigms, concepts, and technologies envelop cloud computing.

Ernemann, C., V. Hamscher, et al. (2004). Benefits of global grid computing for job scheduling. 5th International Workshop on Grid Computing, Pittsburgh, PA, Ieee Computer Soc.

In addition to other advantages, computational Grids are considered to utilize the participating compute resources more efficiently as well as to improve the response time for user jobs. Due to the lack of common large scale global Grids and corresponding studies on Grid workloads this assumption is not yet verified. In this paper, the effect of geographical distribution of Grid resources on the machine utilization and the average response time is analyzed. To this end, simulations have been performed. The results show a significant benefit for the job scheduling quality due to the participation in a true global Grid. The average weighted response times of all submitted jobs decrease up to about 30%. The results have been verified using different workloads and Grid configurations.

Ferguson, J. W. and J. Towns (2001). "The alliance grid." Advances in Engineering Software **32**(5): 417-422.

The advent of high-performance networks available to research institutions has made it possible to transform the way in which computational researchers do work. The goal is to create an environment and supporting infrastructure that allows research and production work on the Grid, as some have named it (Ian Foster, Carl Kesselman, editors. The grid: blueprint for a new computing infrastructure. Los Altos: Morgan Kaufmann, 1999), to be done as easily as working on your desktop. Through the Grid, researchers will have access to resources wherever they are located on the network - resources such as high-performance computers, data archives, scientific instruments, tele-immersive environments, and other limited access assets such as telescopes. The National Computational Science Alliance (Alliance) has the deployment of such a Grid as one of its main missions. (C) 2001 Elsevier Science Ltd. All rights reserved.

Foster, I. and C. Kesselman (1998). The globus project: A status report. 7th Heterogeneous Computing Workshop (HCW 98), Orlando, Fl, Ieee Computer Soc.

The Globus project is a multi-institutional research effort that seeks to enable the construction of computational grids providing pervasive, dependable, and consistent access to high-performance computational resources, despite geographical distribution of both resources and users. Computational grid technology is being viewed as a critical element of future high-performance computing environments that will enable entirely new classes of computation-oriented applications, much as the World Wide Web fostered the development of new classes of information-oriented applications. In this paper, we report on the status of the Globus project as of early 1998. We describe the progress that has been achieved to date in the development of the Globus toolkit, a set of core services for constructing grid tools and applications. We also discuss the Globus Ubiquitous Supercomputing Testbed (GUSTO) that we have constructed to enable large-scale evaluation of Globus technologies, and we review early experiences with the development of large-scale grid applications on the GUSTO testbed.

Foster, I. and C. Kesselman (1999). "The Globus project: a status report." Future Generation Computer Systems **15**(5-6): 607-621.

The Globus project is a multi-institutional research effort that seeks to enable the construction of computational grids providing pervasive, dependable, and consistent

access to high-performance computational resources, despite geographical distribution of both resources and users. Computational grid technology is being viewed as a critical element of future high-performance computing environments that will enable entirely new classes of computation-oriented applications, much as the World Wide Web fostered the development of new classes of information-oriented applications. In this paper, we report on the status of the Globus project as of early 1998. We describe the progress that has been achieved to date in the development of the Globus toolkit, a set of core services for constructing grid tools and applications. We also discuss the Globus Ubiquitous Supercomputing Testbed Organization (GUSTO) that we have constructed to enable large-scale evaluation of Globus technologies, and we review early experiences with the development of large-scale grid applications on the GUSTO testbed. (C) 1998 IEEE.

Foster, I. and C. Kesselman (2000). Computational grids - Invited talk (Reprinted from The Grid: Blueprint for a new computing infrastructure, 1998). 4th International Conference on Vector and Parallel Processing (VECPAR 2000), Oporto, Portugal, Springer-Verlag Berlin.

In this introductory chapter, we lay the groundwork for the rest of the book by providing a more detailed picture of the expected purpose, shape, and architecture of future grid systems. We structure the chapter in terms of six questions that we believe are central to this discussion: Why do we need computational grids? What types of applications will grids be used for? Who will use grids? How will grids be used? What is involved in building a grid? And, what problems must be solved to make grids commonplace? We provide an overview of each of these issues here, referring to subsequent chapters for more detailed discussion.

Foster, I., Y. Zhao, et al. (2008). Cloud Computing and Grid Computing 360-degree compared. Grid Computing Environments Workshop, GCE 2008, Austin, TX.

Cloud Computing has become another buzzword after Web 2.0. However, there are dozens of different definitions for Cloud Computing and there seems to be no consensus on what a Cloud is. On the other hand, Cloud Computing is not a completely new concept; it has intricate connection to the relatively new but thirteen-year established Grid Computing paradigm, and other relevant technologies such as utility computing, cluster computing, and distributed systems in general. This paper strives to compare and contrast Cloud Computing with Grid Computing from various angles and give insights into the essential characteristics of both.

Franke, C., W. Theilmann, et al. (2007). Towards the autonomic business Grid. Proceedings - Fourth IEEE International Workshop on Engineering of Autonomic and Autonomous Systems, EASe 2007, Tucson, AZ.

Business Grids are envisioned to become the adaptive service-oriented utility infrastructure for arbitrary business application satisfying the needs for agility and cost efficiency in today's and tomorrow's ever changing business environment. This industrial report paper discusses the main areas that contribute to this vision, namely Grid computing, autonomic computing and virtualization technologies. Furthermore, it shows how they relate to each other. Â© 2007 IEEE.

Gagliardi, F. (2004). The EGEE European grid infrastructure project. 6th International Conference on High Performance Computing for Computational Science (VECPAR 2005), Valencia, SPAIN, Springer-Verlag Berlin.

The state of computer and networking technology today makes the seamless sharing of computing resources on an international or even global scale conceivable. Extensive computing Grids that integrate large, geographically distributed computer clusters and

data storage facilities have changed from representing a dream to becoming a vision and, with the Enabling Grids for E-science in Europe project (EGEE), a reality today. EGEE aims to provide a European Grid infrastructure for the support of many application domains. This infrastructure is built on the EU Research Network GEANT and exploits Grid expertise that has been generated by previous projects. EGEE is a EU funded project that involves 71 partners from Europe, Russia and the United States. The project started in April 2004 for a first phase of 2 years.

Gagliardi, F., B. Jones, et al. (2005). "Building an infrastructure for scientific Grid computing: status and goals of the EGEE project." Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences **363**(1833): 1729-1742.

The state of computer and networking technology today makes the seamless sharing of computing resources on an international or even global scale conceivable. Scientific computing Grids that integrate large, geographically distributed computer clusters and data storage facilities are being developed in several major projects around the world. This article reviews the status of one of these projects, Enabling Grids for E-SciencE, describing the scientific opportunities that such a Grid can provide, while illustrating the scale and complexity of the challenge involved in establishing a scientific infrastructure of this kind.

Garbacki, P. and V. K. Naik (2007). Efficient resource virtualization and sharing strategies for heterogeneous grid environments. 10th IFIP/IEEE International Symposium on Integrated Network Management 2007, IM '07, Munich.

Resource visualization has emerged as a powerful technique for customized resource provisioning in grid and data center environments. In this paper, we describe efficient strategies for policy-based controlling of virtualization of the physical resources. With these strategies, virtualization is controlled taking into account workload requirements, available capacities of physical resources, and the governing policies. Realizing this control requires simultaneous handling of three problems: (i) determining the virtual resource configurations, (ii) the mapping of resulting virtual resources to physical resources, and (iii) the mapping of workloads to the virtual resources. We pose this as an optimization problem and solve this problem using a linear programming (LP) based approach. We evaluate this approach by implementing it in the Harmony grid environment consisting of heterogeneous resources and heterogeneous workload. Experimental results indicate that our approach is efficient and effective. We extend this approach further by using a two-phase heuristic that allows the decision making component to scale up to handle large scale grid systems. Â© 2007 IEEE.

Gentzsch, W. (2006). D-Grid, an e-science framework for German scientists. 5th International Symposium on Parallel and Distributed Computing, Timisoara, ROMANIA, Ieee Computer Soc.

The World Wide Web has become pervasive in many facets of our professional and personal lives during the past 10 years. Today, another Internet revolution is emerging: Grid Computing. While the Web offers easy access to mostly static information via Hypertext, the Grid adds another fundamental layer by enabling access to and use of the underlying resources. Based on widely accepted grid and web services standards, resources including computers, storage, scientific instruments and experiments, applications, data, and middleware services communicate with each other and deliver results back to the user. These resources are part of a single service-oriented architecture, called OGSA, the Open Grid Services Architecture. For the past several years, early adopters in research and industry have been building and

operating prototypes of grids for global communities, virtual organizations, and within enterprises..

Ghosh, P., N. Roy, et al. (2005). "A pricing strategy for job allocation in mobile grids using a non-cooperative bargaining theory framework." Journal of Parallel and Distributed Computing **65**(11): 1366-1383.

Due to their inherent limitations in computational and battery power, storage and available bandwidth, mobile devices have not yet been widely integrated into grid computing platforms. However, millions of laptops, PDAs and other portable devices remain unused most of the time, and this huge repository of resources can be potentially utilized, leading to what is called a mobile grid environment. In this paper, we propose a game theoretic pricing strategy for efficient job allocation in mobile grids. By drawing upon the Nash bargaining solution, we show how to derive a unified framework for addressing such issues as network efficiency, fairness, utility maximization, and pricing. In particular, we characterize a two-player, non-cooperative, alternating-offer bargaining game between the Wireless Access Point Server and the mobile devices to determine a fair pricing strategy which is then used to effectively allocate jobs to the mobile devices with a goal to maximize the revenue for the grid users. Simulation results show that the proposed job allocation strategy is comparable to other task allocation schemes in terms of the overall system response time. Â© 2005 Elsevier Inc. All rights reserved.

GjermundoÃ¸r, H., M. D. Dikaiakos, et al. (2008). g-Eclipse - An integrated framework to access and maintain grid resources. Proceedings - IEEE/ACM International Workshop on Grid Computing, Tsukuba.

The g-Eclipse framework provides a general, integrated workbench toolset for Grid users, operators and developers. Based on the Open Source Eclipse ecosystem, g-Eclipse supports scientists to interact with Grid resources independent of the underlying Grid middleware. Its main objective is to deliver an extensible framework for different Grid actors, by providing a unified abstraction of the Grid. The Grid abstraction enables Grid application users to access the Grid in a desktop-like manner with wizards specific for common use cases; it also provides a set of visual configuration tools to maintain and configure Grid resources. Â© 2008 IEEE.

Hoffa, C., G. Mehta, et al. (2008). On the Use of Cloud Computing for Scientific Workflows. eScience, 2008. eScience '08. IEEE Fourth International Conference on.

This paper explores the use of cloud computing for scientific workflows, focusing on a widely used astronomy application-Montage. The approach is to evaluate from the point of view of a scientific workflow the tradeoffs between running in a local environment, if such is available, and running in a virtual environment via remote, wide-area network resource access. Our results show that for Montage, a workflow with short job runtimes, the virtual environment can provide good compute time performance but it can suffer from resource scheduling delays and widearea communications.

House, B., P. Marshall, et al. (2008). Grid service hosting on virtual clusters. Proceedings - IEEE/ACM International Workshop on Grid Computing, Tsukuba.

This paper presents an architecture for service hosting on virtual clusters spanning multiple administrative domains that balances the requirements of application developers and resource provider system administrators. The presented architecture and implementation use virtual machines to simplify the deployment of externally-accessible persistent Web and Grid services while allowing resource provider system

administrators to monitor hosted virtual machines and perform critical maintenance when necessary. This approach allows developers full control of their distributed resources, specifies a mechanism for resource provider monitoring and intervention, and reduces the barrier for hosting user-supplied virtual machines on shared resource provider cyberinfrastructure. Â© 2008 IEEE.

Ivanov, II (2007). Utility Computing: Reality and Beyond. 4th International Conference on E-Business and Telecommunication Networks, Barcelona, SPAIN, Springer-Verlag Berlin.

Utility Computing is not a new concept. It involves organizing and providing a wide range of computing-related services as public utilities. Much like water, gas, electricity and telecommunications, the concept of computing as public utility was announced in 1955. Utility Computing remained a concept for near 50 years. Now some models and forms of Utility Computing are emerging such as storage and server virtualization, grid computing, and automated provisioning. Recent trends in Utility Computing as a complex technology involve business procedures that could profoundly transform the nature of companies' IT services, organizational IT strategies and technology infrastructure, and business models. In the ultimate Utility Computing models, organizations will be able to acquire as much IT services as they need, whenever and wherever they need them. Based on networked businesses and new secure online applications, Utility Computing would facilitate "agility-integration" of IT resources and services within and between virtual companies. With the application of Utility Computing there could be concealment of the complexity of IT, reduction of operational expenses, and converting of IT costs to variable 'on-demand' services. How far should technology, business and society go to adopt Utility Computing forms, modes and models?

Jiang, Y., D. Meng, et al. (2005). Adaptive management of a utility computing. IEEE International Conference on Cluster Computing, Burlington, MA, Ieee.

Kacsuk, P., A. Kertesz, et al. (2008). Can We Connect Existing Production Grids into a World Wide Grid? 8th International Conference on High Performance Computing for Computational Science (VECPAR 2008), Toulouse, FRANCE, Springer-Verlag Berlin.

The World Wide Web has become a phenomenon that now influences our everyday life in any possible areas and disciplines. This paper investigates how a grid equivalent of the WWW, the World Wide Grid can be created. We define requirements towards a work flow-oriented computational World Wide Grid and propose a solution how current production Grids call be connected in order to form the technical basis of this infrastructure. A meta-broker concept and its utilization to achieve the highest level of parallelism by the created architecture in a user transparent way are explained.

Kawato, M., F. Machida, et al. (2006). Service-aware virtual machine placement mechanism for utility computing. 18th IASTED International Conference on Parallel and Distributed Computing and Systems, Dallas, TX, Acta Press Anaheim.

Dynamic server resource allocation to services on networks, or utility computing, is a powerful technology to provide required computer resources for multiple service providers at low cost. Virtual machine (VM) technology can be combined with utility computing to further improve server resource utilization. An important technical issue about VM-based utility computing is optimal placement of VMs on physical server nodes, because performance of services may be seriously affected by VM placement. We address this issue by introducing an on-line placement algorithm on the basis of performance influence among services. We have implemented this algorithm and evaluated on a simulated environment. The results have shown that the proposed

mechanism can get roughly 25% better score calculated by the placement rules compared with a random placement algorithm.

Komisarczuk, P. and K. Bubendorfer (2005). A position paper: Towards an utility computing and communications infrastructure. Proceedings of the 2005 International Conference on Communications in Computing, CIC'05, Las Vegas, NV.

The future of the Internet may be a conjunction of Grid computing and service overlays hosted over Next Generation Internet (NGI) technologies, consisting of both wireline and wireless networks. The Grid and overlay networks can be seen as a key service layer for the future where the combination of computing and communication resources is dynamically allocated to virtual organizations on demand to enable optimal service and business deployment. To enable this synergy the efficient allocation of resources at minimum cost needs to be enabled. Today our resource description, resource matching and pricing in Grid/NGI needs to be extended/addressed and our viewpoint on these aspects are introduced in this paper. Initially this paper looks at some of the drivers and business models that could be envisaged with some background on developing grid networks and a brief review of some recent work on QoS description.

Laccetti, G. and G. Schmid (2007). "A framework model for grid security." Future Generation Computer Systems-the International Journal of Grid Computing Theory Methods and Applications **23**(5): 702-713.

Computational grids can be considered as tiered objects; following a widespread terminology, cluster grids may be grouped into enterprise grids, that in turn may belong to global grids. Therefore, computing grid security has to be "tiered" too, with the ground level remaining the OS one. In this work, we introduce a sort of unified approach, an overall architectural framework for access control to grid resources, and one that adheres as much as possible to current security principles. Current grid security implementations are viewed in the light of this model, their main drawbacks are described, and we show how our proposal is able to avoid them. We believe that a main strategy could be to adopt both PKI and PMI infrastructures at the grid layer, ensuring that an adequate transfer of authentication and authorization will be made between the Virtual Organization and Resource Provider layers. This can be achieved by extending those features at the OS layer as system applications and services. (c) 2007 Elsevier B.V. All rights reserved.

Lana, G. Q. and C. B. Westphall (2008). User Maturity Based Trust Management for Grid Computing. Networking, 2008. ICN 2008. Seventh International Conference on.

Trust management has been considered an important factor in Grid computing security. Grids require a secure way to establish trust in their participants without requiring continuous intervention by the system administrators. This paper describes trust management in Grid computing systems and its use in evaluation environments to update the trust levels. This approach can be incorporated in Grid security systems to assist in issuing authorisations, in resource and service management, and in decision-making processes.

Leinberger, W. and V. Kumar (1999). "Information Power Grid: The new frontier in parallel computing?" Ieee Concurrency **7**(4): 75-84.

Li, C. and L. Li (2007). Utility based multiple QoS guaranteed resource scheduling optimization in grid computing. Proceedings - International Conference on Computing: Theory and Applications, ICCTA 2007, Kolkata.

In this paper, we consider multiple QoS based grid resource scheduling. It is heterogeneity and dynamics of the grid that make QoS problems in grid environment challenging. Computational grid's resource management must deal with various demands from users. Each of grid task agent's diverse requirements is modeled as a quality of service (QoS) dimension, associated with each QoS dimension is a utility function that defines the benefit that is perceived by a user with respect to QoS choices in that dimension. The objective of multiple QoS based grid resource scheduling is to maximize the global utility of the scheduling system. This paper proposes an iterative scheduling algorithm that is used to perform optimal multiple QoS based resource scheduling. The experiments show that optimal multiple QoS based resource scheduling involves less overhead and leads to more efficient resource allocation than no optimal resource allocation. © 2007 IEEE.

Li, C. L. and L. Y. Li (2007). "An optimization approach for decentralized QoS-based scheduling based on utility and pricing in Grid computing." Concurrency and Computation-Practice & Experience **19**(1): 107-128.

This paper presents an optimization approach for decentralized Quality of Service (QoS)-based scheduling based on utility and pricing in Grid computing. The paper assumes that the quality dimensions can be easily formulated as utility functions to express quality preferences for each task agent. The utility values are calculated by the user-supplied utility function that can be formulated with the task parameters. The QoS constraint Grid resource scheduling problem is formulated into a utility optimization problem. The QoS-based Grid resource scheduling optimization is decomposed into two subproblems by applying the Lagrangian method. In the Grid, a Grid task agent acts as a consumer paying for the Grid resource and the resource providers receive profits from task agents. A pricing-based QoS scheduling algorithm is used to perform optimally decentralized QoS-based resource scheduling. The experiments investigate the effect of the QoS metrics on the global utility and compare the performance of the proposed algorithm with other economical Grid resource scheduling algorithms. Copyright (c) 2006 John Wiley & Sons, Ltd.

Li, L., W. Yuanzhuo, et al. (2005). Utility-based on-demand heuristic strategy to grid computing. Proceedings of the Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services, ICAS/ICNS 2005, Papeete, Tahiti.

Grids aim at exploiting synergies that result from cooperation of autonomous distributed entities. For this cooperation to be sustainable, participants need to have economic incentive. Within service-oriented Grids, resources may be advertised and traded as services, and QoS requirement is the most important attribute. In this paper, we build on two intellectual threads: market-based approaches and the economic theory of utility to resource allocation. We consider the problem of task scheduling subject to time and budget constrains as QoS-aware joint utility function, and scheduling object is to maximize total attained utility. This scheduling problem can be shown to be NP-hard, thus we present a utility-driven heuristic scheduling algorithm (UDS) for this problem. We study the performance of UDS algorithm through a Stochastic Petri Net (SPN) model. In the end, the performance of model is analyzed by SPNP software package. The numerical results show that UDS can guarantee QoS for users. Â© 2005 IEEE.

Liang-Jie, Z. and L. Zongwei (2005). "Grid and utility computing promises a bright future." International Journal of Grid and Utility Computing **1**(1): 1-3.

A bright future is promised ahead for grid computing, a technology expected to become a popular enabling means for business applications. Today, most business

applications are developed on top of systems depending on costly `islands of computation', with each application needing its own dedicated machines and software. Grid computing breaks the barries and bridges the gaps by bringing together `resources islands', forming virtualized computing resources. Further, it opens the door to utility computing to provide electric utility like services. This service orientation shift has resulted in new challenges and opportunities for grid research and development

Litke, A., K. Konstanteli, et al. (2008). "Managing service level agreement contracts in OGSA-based Grids." Future Generation Computer Systems **24**(4): 245-258.

Grids and mobile Grids can form the basis and the enabling technology for pervasive and utility computing due to their ability to be open, highly heterogeneous and scalable. However, the process of selecting the appropriate resources and initiating the execution of a job is not enough to provide quality in a dynamic environment such as a mobile Grid, where changes are numerous, highly variable and with unpredictable effects. In this paper we present a scheme for advancing and managing Quality of Service (QoS) attributes contained in Service Level Agreement (SLA) contracts of Grids that follow the Open Grid Services Architecture (OGSA). In order to achieve this, the execution environment of the Grid infrastructure establishes and exploits the synergies between the various modules of the architecture that participate in the management of the execution and the enforcement of the SLA contractual terms. We introduce an Execution Management Service which is in collaboration with both the application services and the network services in order to provide an adjustable quality of the requested services. The components that manage and control the execution in the Grid environment interact with the suit of the SLA-related services exchanging information that is used to provide the quality framework of the execution with respect to the agreed contractual terms. The described scheme has been implemented in the framework of the Akogrimo IST project. Â© 2007 Elsevier Ltd. All rights reserved.

Liu, L., Y. Wang, et al. (2006). Utility-based computing model for grid. Proceedings - First International Conference on Semantics, Knowledge and Grid, SKG 2005, Beijing.

Within service-oriented Grids, resources may be advertised and traded as services and participants need to have economic incentive. Grid can be seen as a multi-agent system, agent are rational and state their QoS requirements in the form of a utility function which maximize is his goal. In this paper, we propose utility-based computing model for grid. We consider the problem of task scheduling subject to time and budget constrains as joint utility function, and scheduling object is to maximize total attained utility. This scheduling problem can be shown to be NP-hard, thus we present a joint utility function scheduling algorithm (JUFS) for this problem. The performance of model is analyzed by SPNP software package. The numerical results show that JUFS can provide user more utility. © 2006 IEEE.

Llorente, I. M., R. S. Montero, et al. (2006). A grid infrastructure for utility computing. Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WET ICE, Manchester.

Utility computing is a service provisioning model, which will provide adaptive, flexible and simple access to computing resources, enabling a pay-per-use model for computing similar to traditional utilities such as water or electricity. The deployment of a utility computing solution involves a full separation between the provider and the consumer. The consumer requires a uniform, secure and reliable functionality to access the utility computing service and the provider requires a scalable, flexible and adaptive infrastructure to provide the service. The solution should be based on

standards and allow a gradual deployment in order to obtain a favourable response from the application developers and the information technology staff. Grid technology overcomes such challenges by means of its standard functionality for flexible integration of diverse distributed resources. This position paper proposes an innovative solution for utility computing which can be deployed on a Grid infrastructure based on Globus toolkit and GridWay components Â© 2006 IEEE.

Losup, A., D. H. J. Epema, et al. (2007). Inter-operating grids through delegated matchmaking. Proceedings of the 2007 ACM/IEEE Conference on Supercomputing, SC'07, Reno, NV.

The grid vision of a single computing utility has yet to materialize: while many grids with thousands of processors each exist, most work in isolation. An important obstacle for the effective and efficient inter-operation of grids is the problem of resource selection. In this paper we propose a solution to this problem that combines the hierarchical and decentralized approaches for interconnecting grids. In our solution, a hierarchy of grid sites is augmented with peer-to-peer connections between sites under the same administrative control. To operate this architecture, we employ the key concept of delegated matchmaking, which temporarily binds resources from remote sites to the local environment. With trace-based simulations we evaluate our solution under various infrastructural and load conditions, and we show that it outperforms other approaches to inter-operating grids. Specifically, we show that delegated matchmaking achieves up to 60% more goodput and completes 26% more jobs than its best alternative. (c) 2007 ACM.

Machiraju, V., A. Sahai, et al. (2003). Managed utility computing: The grid as management backplane. 1st Latin-American Symposium on Dependable Computing, Sao Paulo, Brazil, Springer-Verlag Berlin.

Enterprise IT exhibits increasingly complex networked systems and distributed applications, making the task of an IT operator or administrator exceedingly difficult. We argue in the keynote associated with this paper for the necessity of a clean, standardized, service-centric software architecture to automate and facilitate operator tasks throughout the life-cycle of systems and applications. We refer to the proposed solution as 'managed utility computing,' since it enables utility computing while improving manageability. The proposed architecture is a web services based 'grid' architecture, with targeted management extensions. This short note argues that this architecture is also necessary and appropriate as backplane for traditional management software (irrespective of the utility computing context), to support increasingly complex management tasks for increasingly complex systems.

Maheswaran, M., B. Maniymaran, et al. (2004). Towards a quality of service aware public computing utility. 3rd IEEE International Symposium on Network Computing and Applications, Boston, MA, Ieee Computer Soc.

This paper describes a design for a quality of service aware public computing utility (PCU). The goal of the PCU is to utilize the idle capacity of the shared public resources and augment the capacity with dedicated resources as necessary, to provide high quality of service to the clients at the least cost. Our PCU design combines peer-to-peer (P2P) and Grid computing ideas in a novel manner to construct a utility-based computing environment. In this paper we present the overall architecture and describe two major components: a P2P overlay substrate for connecting the resources in a global network and a community-based decentralized resource management system.

Maniymaran, B. and M. Maheswaran (2004). Virtual clusters: A dynamic resource coallocation strategy for computing utilities. Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Systems, Cambridge, MA.

This paper presents a concept called virtual clusters (VCs) to allocate resources for an application from a computing utility with a geographically distributed resource base. The VC creation process is modeled as a facility location problem and an efficient heuristic is devised to solve it. We extend the model to include an "overload partition" to a VC such that demand surges can be efficiently handled. Extensive simulations have been conducted to examine the performance of VCs under different scenarios and to compare it with a fully dynamic scheme called the Service Grid. The results indicate that VC is more cost-effective and robust than Service Grid.

Maoz, T., A. Barak, et al. (2008). Combining virtual machine migration with process migration for HPC on multi-clusters and grids. Proceedings - IEEE International Conference on Cluster Computing, ICCC, Tsukuba.

The renewed interest in virtualization gives rise to new opportunities for running High Performance Computing (HPC) applications on clusters and Grids. These include the ability to create a uniform (virtual) run-time environment on top of a multitude of hardware and software platforms, and the possibility for dynamic resource allocation towards the improvement of process performance, e.g., by Virtual Machine (VM) migration as a means for load-balancing. This paper deals with issues related to running HPC applications on multi-clusters and Grids using VMware, a virtualization package running on Windows, Linux and OS X. The paper presents the "Jobrun" system for transparent, on-demand VM launching upon job submission, and its integration with the MOSIX cluster and Grid management system. We present a novel approach to job migration, combining VM migration with process migration using Jobrun, by which it is possible to migrate groups of processes and parallel jobs among different clusters in a multi-cluster or in a Grid. We use four real HPC applications to evaluate the overheads of VMware (both on Linux and Windows), the MOSIX cluster extensions and their combination, and present detailed measurements of the performance of Jobrun. Â© 2008 IEEE.

Medernach, E. (2005). Workload analysis of a cluster in a grid environment. 11th International Workshop Job Scheduling Strategies for Parallel Processing, Cambridge, MA, Springer-Verlag Berlin.

With Grids, we are able to share computing resources and to provide for scientific communities a global transparent access to local facilities. In such an environment the problems of fair resource sharing and best usage arise. In this paper, the analysis of the LPC cluster usage (Laboratoire de Physique Corpusculaire, Clermont-Ferrand, France) in the EGEE Grid environment is done, and from the results a model for job arrival is proposed.

Menasce, D. A. and E. Casalicchio (2004). "QoS in grid computing." Internet Computing, IEEE **8**(4): 85-87.

Grid computing is already a mainstream paradigm for resource-intensive scientific applications, but it also promises to become the future model for enterprise applications. The grid enables resource sharing and dynamic allocation of computational resources, thus increasing access to distributed data, promoting operational flexibility and collaboration, and allowing service providers to scale efficiently to meet variable demands. Large-scale grids are complex systems composed of thousands of components from disjoined domains. Planning the capacity to guarantee quality of service (QoS) in such environments is a challenge because global

service-level agreements (SLAs) depend on local SLAs. We provide a motivating example for grid computing in an enterprise environment and then discuss the how resource allocation affects SLAs.

Montes, J., A. Sanchez, et al. (2008). The Grid as a Single Entity: Towards a Behavior Model of the Whole Grid. On the Move Confederated International Conference and Workshops, Monterrey, MEXICO, Springer-Verlag Berlin.

Grids emerged in the last decade as large distributed environments where the greatest challenges of the scientific community could be faced. These challenges, commonly known as grand challenge applications (GCAs), are characterized as problems that, given their size and/or complexity, can not be solved by means of conventional computing techniques. The grid scalability, flexibility and massive computation power make it an ideal environment to face these GCAs. In order to provide the expected service and achieve the required performance for these applications, grids must have powerful management mechanisms that efficiently deal with the natural variability and heterogeneity of the environment. Since the birth of this technology, this has been one of the key aspects of its development and, in many ways, one of the most problematic. Conventional grid management mechanisms try to improve performance based on the individual analysis of every component on the system. Then they intend to adjust the configuration or predict the behavior of each independent element. This approach may seem reasonable considering the large scale and complexity of the grid, but could possibly fail to achieve optimal performance, because in most cases it lacks the capability to understand the effects that different elements have on each other when they work together. From a more theoretical point of view, if we consider "the grid" as an individual entity (with its computational power, storage capacity and so on), it seems logical to analyze it as such, instead of composed of a huge set of individual resources. This could be similar as how computers are regarded as individual entities, even though they are made of several electronic components of different nature, or how clusters are most times considered as single machines, when in fact they are composed by many computers. In this paper we present a set of techniques oriented to build a global behavioral model of a large scale distributed environment, such as a grid. This kind of model would provide a deep understanding of the system and a better reference for management and optimization. A case Study is also presented, in order to provide a validation of this approach.

Narmadha, K. and S. T. Selvi (2007). A three level broker framework for global service - Oriented utility grid. ICICT 2007: Proceedings of International Conference on Information and Communication Technology, Dhaka.

The topology of brokers is very important in a global service-oriented utility grid for serving both the users' and service providers' interests. A three level broker framework is proposed in this paper in which the brokers are not bound to any particular domain. Hence the users and service providers are free to choose the broker they wish. The framework is designed in such a way that the brokers are organized into three levels. A Level 1(L1) broker is assisted by many Level 2(L2) brokers and a L2 broker may be assisted by many Level 3(L3) brokers. Users interact with L1 brokers only. L2 brokers are classified according to the type of resources. L3 brokers are classified based on the resources under a L2 broker. A L1 broker contacts L2 brokers, which in turn may contact L3 brokers for selecting the requested resources.

Obozinski, V. R. (2005). From grid technologies (in 6th of IST RTD 2002-2006) to knowledge utility. ISPDC 2005: 4th International Symposium on Parallel and Distributed Computing, Lille.

An overview of the Strategic Objective on Advanced Grid Technologies, Systems and Services under the IST (Information Society Technologies) thematic priority of the 6th European RTD Framework Programme (2002-2006) will be provided. The evolution and progress since the 5th Framework Programme will be sketched. The presentation will include a brief description of the portfolio of 12 new projects - with a total EU funding of 52 million ?- that were launched in the 2nd half of 2004. Latest strategic planning activities with regard to the Strategic Objective -particularly in the context of the European Research Area - will be covered in relative details. The related topics (S.O. 2.5.4) of the IST Work Programme 2005-2006 - as specified for Call 5 - will be addressed in an exhaustive manner. Future perspectives for the 7th Framework Programme (2007-2013) regarding the ICT technology pillar on/ Software, Grids, Security and Dependability/ as laid out in the Commission proposal of April 2005 will be outlined on a high level. A question and answer time will follow the presentation. © 2005 IEEE.

Ogawa, H., S. Itoh, et al. (2007). "GridASP: an ASP framework for grid utility computing." Concurrency and Computation Practice & Experience **19**(6): 885-91.

One of the greatest evolutions brought about by grid technology is 'grid utility computing', which utilizes various kinds of IT resources and applications across multiple organizations and enterprises, and integrates them into a comprehensive and valuable service. Since 2004, we proposed and have been developing the GridASP framework, which realizes grid-enabled application service providers (ASPs) in order to realize grid utility computing. GridASP can bind application providers, resource providers, and service providers together and provide application execution services with security and anonymity to enterprise/science users. In this article, we report the conceptual idea of GridASP and the detail of the framework now being developed. Information on GridASP can be found at http://www.gridasp.org.

Olugbara, O. O., M. O. Adigun, et al. (2007). Utility grid computing and body area network as enabler for ubiquitous rural e-healthcare service provisioning, Piscataway, NJ, USA, IEEE.

Body area network (BAN) has the potential to revolutionalize healthcare management by allowing the use of innovative smart sensors for remote diagnosis, treatment and continuous monitoring of a patient's body conditions. However, there are many challenges such as low bandwidth, limited computational power and limited functionality for communication with wireless sensor networks. On the other hand, utility grid computing (UGC) is rich in functionality, can be used to provide access to appropriate computational services and can give high-bandwidth to a large collection of distributed time-varying resources. This paper describes an integrated UGC-BAN technology as an enabler for ubiquitous healthcare service provisioning. This would help addressing the challenges of accessibility to quality healthcare services and costs of provisioning to rural communities of developing countries.

Opitz, A., H. KÃ¶nig, et al. (2008). "What does grid computing cost?" Journal of Grid Computing **6**(4): 385-397.

Grid computing has gained considerable attention in research and industry. High expectations are associated with the approach. However, so far only few papers have been published about the costs caused by Grid computing. In this article we pursue two main goals: to analyze the different types of costs and to determine the total costs of a resource provider. Our approach is based on the discussion of general cost categories that have to be taken into account. We give concrete numbers for the different categories and use these numbers to estimate the costs in two real life Grids:

the EGEE project and the Grid of the pharmaceutical company Novartis. A summarizing discussion concludes the paper. Â© Springer Science+Business Media B.V. 2008.

Ouyang, J. S., A. Sahai, et al. (2007). A mechanism of specifying and determining pricing in utility computing environments. 2nd IEEE/IFIP International Workshop on Business-Driven IT Management, Munich, GERMANY, Ieee.

Utility services providers offer on-demand services for multiple users and applications from a shared resource pool. The utility computing environment provider manages resources that have multiple dimensions, namely CPU, Memory and Disk. It is important that a mechanism of quantification he utilized for determining pricing. Also in these environments, we believe that price needs to be adapted to reflect and influence dynamic change in demand on resources, and can be used, as an economic mechanism, to efficiently manage a utility computing environment. To achieve this, we propose in this paper an adaptive pricing mechanism. The objective of this scheme is twofold: it, for business objectives, adapts price based upon the supply and demand curves over time intervals; on the other hand, price invisibly manages the distribution of resource demand across time intervals, therefore makes resource utilization more efficient and balanced over time. In this approach, the resource unit to be priced is called capacity unit, and applications are billed based upon the capacity units that have been reserved or used.

Palankar, M., A. Lamnitchi, et al. (2008). Amazon S3 for science grids: A viable solution? International Symposium on High Performance Distributed Computing, HPDC 2008 - Proceedings of the 2008 International Workshop on Data-aware Distributed Computing 2008, DADC'08, Boston, MA.

Amazon.com has introduced the Simple Storage Service (S3), a commodity-priced storage utility. S3 aims to provide storage as a low-cost, highly available service, with a simple 'pay-as-you-go' charging model. This article makes three contributions. First, we evaluate S3's ability to provide storage support to large-scale science projects from a cost, availability, and performance perspective. Second, we identify a set of additional functionalities that storage services targeting data-intensive science applications should support. Third, we propose unbundling the success metrics for storage utility performance as a solution, to reduce storage costs. Copyright 2008 ACM.

Pandey, S., J. Chao, et al. (2008). Gridbus Work?ow Management System on Clouds and Global Grids. eScience, 2008. eScience '08. IEEE Fourth International Conference on.

The Gridbus workflow management system (GWMS) is designed to execute scientific applications, expressed in the form of workflows, onto Grid and Cloud resources. With the help of this system, we demonstrate a computational and data-intensive Image Registration (IR) workflow for functional Magnetic Resonance Imaging (fMRI) applications on Clouds and global Grids. We also present a demonstration of the Aneka System. Aneka is a .NET based Cloud software platform that provides: (i) a configurable service container hosting pluggable services for discovering, scheduling various types of workloads and (ii) a flexible and extensible framework supporting various programming models. We use distributed Grid and Cloud resources from Australia, Austria, France, Japan, and USA for the executions of image registration workflow, and resources from Melbourne University for executing map-reduce applications in Aneka cloud environment.

Peng, L., L. K. Ng, et al. (2005). YellowRiver: A flexible high performance cluster computing service for grid, Beijing, China, Inst. of Elec. and Elec. Eng. Computer Society.

Computational Grids provide an emerging highly distributed computing platform for scientific computing. Recently, service oriented architecture (SOA) has become a trend of implementing software systems including Grids. SOA provides more flexibilities for Grid users at the service level. High performance computing (HPC) facilities such as HPC clusters, as building blocks of Grid computing, are playing an important role in computational Grid and they are embracing SOA when integrated into Grid in the format of services. Currently, how to build flexible and easy-to-use HPC service for Grid computing still remains an open topic and not much work has been done in this area. In this paper, we propose an HPC cluster service architecture for Grid computing and utility computing. It provides the basic function such as service deployment, service monitoring, and service execution, etc. HPC cluster service deployment not only includes normal application deployment, but also operating system (currently Open Solaris) deployment on demand. Based on Solaris Jumpstart technology and some related tools, a prototype of this architecture has been developed and running on HPC clusters. With our prototype, the Grid users are able to deploy a basic HPC environment (e.g., OpenSolaris, MPICH, Sun Grid Engine or N1 Grid Engine resource management tool) among the available cluster nodes. Our experiments show that our work provide great convenience and flexibility for users to setup and customize their preferred HPC cluster environment for their computation intensive applications in Grid computing or utility computing. 2005 IEEE.

Pugliese, A., D. Talia, et al. (2008). "Modeling and supporting grid scheduling." Journal of Grid Computing **6**(2): 195-213.

Grid resource management systems and schedulers are important components for building Grids. They are responsible for the selection and allocation of Grid resources to current and future applications. Thus, they are important building blocks for making Grids available to user communities. In this paper we briefly analyze the requirements of Grid resource management and provide a classification of schedulers. Then, we define an extensible formal model for Grid scheduling activities, and characterize the general Grid scheduling problem. Finally, we provide a reference architecture for the support of our model and discuss different aspects of architectural implementations. © Springer Science + Business Media B.V. 2007.

Quétier, B., V. Neri, et al. (2007). "Scalability comparison of four host virtualization tools." Journal of Grid Computing **5**(1): 83-98.

Virtualization tools are becoming popular in the context of Grid Computing because they allow running multiple operating systems on a single host and provide a confined execution environment. In several Grid projects, virtualization tools are envisioned to run many virtual machines per host. This immediately raises the issue of virtualization scalability. In this paper, we compare the scalability merits of Four virtualization tools. First, from a simple experiment, we motivate the need for simple microbenchmarks. Second, we present a set of metrics and related methodologies. We propose four microbenchmarks to measure the different scalability parameters for the different machine resources (CPU, memory disk and network) on three scalability metrics (overhead, linearity and isolation). Third, we compare four virtual machine technologies (Vserver, Xen, UML and VMware). The results of this study demonstrate that all the compared tools present different behaviors with respect to scalability, in terms of overhead, resource occupation and isolation. Thus this work will help user to select tools according to their application characteristics. © Springer Science + Business Media B.V. 2006.

Rahman, M. and R. Buyya (2008). An autonomic workflow management system for global Grids. Proceedings CCGRID 2008 - 8th IEEE International Symposium on Cluster Computing and the Grid, Lyon.

Workflow Management System is generally utilized to define, manage and execute workflow applications on Grid resources. However, the increasing scale complexity, heterogeneity and dynamism of Grid environment that includes networks, resources and applications have made such workflow management systems brittle, unmanageable and insecure. Autonomic computing provides a holistic approach for the design and development of systems/applications that can adapt themselves to meet requirements of performance, fault tolerance, reliability, security, etc., without manual intervention. Therefore, this research aims to design and develop mechanisms for building an autonomic workflow management system that will incorporate the properties of autonomic computing and exhibit the ability to reconfigure itself to the changes in the Grid environment, discover, diagnose and react to the disruptions of workflow execution, and monitor and tune Grid resources automatically. © 2008 IEEE.

Reed, D. A. (2008). Clouds, clusters and ManyCore: The revolution ahead. Cluster Computing, 2008 IEEE International Conference on.

Without doubt, scientific discovery, business practice and social interactions are moving rapidly from a world of homogeneous and local systems to a world of distributed software, virtual organizations and cloud computing infrastructure, all powered by multicore processors and large-scale infrastructure. In science, a tsunami of new experimental and computational data and a suite of increasingly ubiquitous sensors pose vexing problems in data analysis, transport, visualization and collaboration. In society and business, software as a service and cloud computing are empowering distributed groups. Letpsilas step back and think about the longer term future. Where is the technology going and what are the implications? What architectures are appropriate? How to we manage power and scale? What are the right size building blocks? How do we come to grips with the fact that our clusters and data centers are now bigger than the Internet was just a few years ago? How do we develop and support malleable software? What is the ecosystem of components in which distributed, data rich applications will operate? How do we optimize performance and reliability? How do we program these systems?

Richard, R. J. A., A. A. Joshi, et al. (2008). "Implementation of computational grid services in enterprise grid environments." American Journal of Applied Sciences **5**(11): 1442-1447.

Grid Computing refers to the development of high performance computing environment or virtual super computing environment by utilizing available computing resources in a LAN, WAN and Internet. This new emerging research field offers enormous opportunities for e-Science applications such as astrophysics, bioinformatics, aerospace modeling, cancer research etc. Grid involves coordinating and sharing of computing power, application, data storage, network resources etc., across dynamically and geographically dispersed organizations. Most Grid environments are developed using Globus toolkit which is a UNIX/Linux based middleware to integrate computational resources over the network. The emergence of Global Grid concept provides an excellent opportunity for Grid based e-Science applications to use high performance super computing environments. Thus windows based enterprise grid environments can't be neglected in the development of Global Grids. This study discusses the basics of enterprise grids and the implementation of enterprise computational grids using Alchemi Tool Kit. This review study is organized into three parts. They are (i) Introduction of Grid Technologies, (ii) Design Concepts of

Enterprise Grids and (iii) Implementation of Computational Grid Services. © 2008 Science Publications.

Sahai, A., S. Singhal, et al. (2004). Automated policy-based resource construction in utility computing environments. 9th IEEE/IFIP Network Operations and Management Symposium (NOMS 2004), Seoul, SOUTH KOREA, Ieee.

A utility environment is dynamic in nature. It has to deal with a large number of resources of varied types, as well as multiple combinations of those resources. By embedding operator and user level policies in resource models, specifications of composite resources may be automatically generated to meet these multiple and varied requirements. This paper describes a model for automated policy-based construction of complex environments. We pose the policy problem as a goal satisfaction problem that can be addressed using a constraint satisfaction formulation. We show how a variety of construction policies can be accommodated by the resource models during resource composition. We are implementing this model in a prototype that uses CIM as the underlying resource model and exploring issues that arise as a result of that implementation.

Samimi, F. A., P. K. McKinley, et al. (2007). "Service Clouds: Distributed Infrastructure for Adaptive Communication Services." Network and Service Management, IEEE Transactions on **4**(2): 84-95.

This paper describes service clouds, a distributed infrastructure designed to facilitate rapid prototyping and deployment of adaptive communication services. The infrastructure combines adaptive middleware functionality with an overlay network substrate in order to support dynamic instantiation and reconfiguration of services. The service clouds architecture includes a collection of low-level facilities that can be invoked directly by applications or used to compose more complex services. After describing the service clouds architecture, we present results of experimental case studies conducted on the PlanetLab Internet testbed alone and a mobile computing testbed.

Sandholm, T., K. Lai, et al. (2006). Market-based resource allocation using price prediction in a high performance computing grid for scientific applications. Proceedings of the IEEE International Symposium on High Performance Distributed Computing, Paris.

We present the implementation and analysis of a market-based resource allocation system for computational Grids. Although Grids provide a way to share resources and take advantage of statistical multiplexing, a variety of challenges remain. One is the economically efficient allocation of resources to users from disparate organizations who have their own and sometimes conflicting requirements for both the quantity and quality of services. Another is secure and scalable authorization despite rapidly changing allocations. Our solution to both of these challenges is to use a market-based resource allocation system. This system allows users to express diverse quantity- and quality-ofservice requirements, yet prevents them from denying service to other users. It does this by providing tools to the user to predict and tradeoff risk and expected return in the computational market. In addition, the system enables secure and scalable authorization by using signed moneytransfer tokens instead of identity-based authorization. This removes the overhead of maintaining and updating access control lists, while restricting usage based on the amount of money transferred. We examine the performance of the system by running a bioinformatics application on a fully operational implementation of an integrated Grid market. Â© 2006 IEEE.

Satayapiwat, C., R. Egawa, et al. (2008). A utility-based double auction mechanism for efficient Grid resource allocation. Proceedings of the 2008 International Symposium on Parallel and Distributed Processing with Applications, ISPA 2008, Sydney, NSW.

In Grid Computing, harnessing the power of idle resources in a distributed environment is one of the important features. However, to fully benefit from this computing model, an appropriated resource allocation method needs to be carefully chosen and deployed. A number of studies have been done on this area and one of the promising approaches is to adopt a marketing scheme called the auction model, which has been drawing much attention during past several years. In this paper, we propose a new utility-aware resource allocation protocol to make external scheduling decision in Grid. Users and service providers specify one or more weight values, and then, an auctioneer uses these values for calculating both users' and service providers' utility values which reflect preference upon the matched members in the different group. Then, we map the scheduling problem with these utility values into the problem in a weighted bipartite graph, and propose a new matching algorithm based on the existing SMP (Stable Marriage Problem) matching algorithm. Finally, the performance of this auction's awarding technique is evaluated. Â© 2008 IEEE.

Shaheen, G., M. U. Malik, et al. (2006). Grid visualizer: A monitoring tool for grid environment. Proceedings - International Workshop on Database and Expert Systems Applications, DEXA, Copenhagen.

One specific problem in wide-area distributed computing environment is effective management of the vast amount of resources that are made available within the grid environment. This paper discusses bandwidth partitioning, fault tolerance, network congestion and scalability issues of grid monitoring and proposes implementation of Dynamic Bandwidth Partitioning algorithm (Q-DBP), on-demand execution of events and agents based monitoring approach respectively to solve these issues. The architecture integrates with the grid middleware layer. This agent infrastructure provides support for mobile agents that are scalable in number and in number of resources. © 2005 IEEE.

Smith, M., M. Engel, et al. (2006). Security issues in on-demand grid and cluster computing. Sixth IEEE International Symposium on Cluster Computing and the Grid Workshops, 2006. CCGRID 06.

In this paper, security issues in on-demand Grid and cluster computing are analyzed, a corresponding threat model is presented and the challenges with respect to authentication, authorization, delegation and single sign-on, secure communication, auditing, safety, and confidentiality are discussed. Three different levels of on-demand computing are identified, based on the number of resource providers, solution producers and users, and the trust relationships between them. It is argued that the threats associated with the first two levels can be handled by employing operating system virtualization technologies based on Xen, whereas the threats of the third level require the use of hardware security modules proposed in the context of the Trusted Computing Platform Alliance (TCPA). The presented security mechanisms increase the resilience of the service hosting environment against both malicious attacks and erroneous code. Thus, our proposal paves the way for large scale hosting of Grid or web services in commercial scenarios. © 2006 IEEE.

Song, Y., Y. Q. Li, et al. (2008). A Service-Oriented Priority-Based Resource Scheduling Scheme for Virtualized Utility Computing. 15th International Conference on High Performance Computing (HiPC 2008), Bangalore, INDIA, Springer-Verlag Berlin.

In order to provide high resource utilization and QoS assurance in utility computing hosting concurrently various services, this paper proposes a service computing framework-RAINBOW for VM(Virtual Machine)-based utility computing. In RAINBOW, we present a priority-based resource scheduling scheme including resource flowing algorithms (RFaVM) to optimize resource allocations amongst services. The principle of RFaVM is preferentially ensuring performance of some critical services by degrading of others to some extent when resource competition arises. Based on our prototype, we evaluate RAINBOW and RFaVM. The experimental results show that RAINBOW without RFaVM provides 28%similar to 324% improvements in service performance, and 26% higher the average CPU utilization than traditional service computing framework (TSF) in typical enterprise environment. RAINBOW with RFaVM further improves performance by 25%similar to 42% for those critical services while only introducing up to 7% performance degradation to others, with 2%similar to 8% more improvements in resource utilization than RAINBOW without RFaVM.

Song, Y., Y. Z. Sun, et al. (2007). An adaptive resource flowing scheme amongst VMs in a VM-based utility computing. 7th IEEE International Conference on Computer and Information Technology, Aizu-Wakamatsu City, JAPAN, Ieee Computer Soc.

In order to optimize the using of server resources which host different services such as web services, this paper describes an adaptive and dynamic resource flowing scheme amongst VMs in a VM-based utility computing environment called ADVM (Adaptive and Dynamic Virtual Machine). In our scheme, VMs adjust their resources (CPU and memory) adaptively to share the physical resources efficiently. ADVM provides an on-the-fly, transparent and lazy resource flowing amongst VMs over a single server. By using dynamic priority and the strategy of adjusting resources, no VM interferes with others even via contention for some resources. To demonstrate the elegance of our strategy, we present a prototype implementation of AD VM. The experimental results indicate that AD VM effectively enforce the system performance based on Xen.

Soni, D. and J. Sharma (2007). Role of grid computing in Indian education. 12th WSEAS International Conference on APPLIED MATHEMATICS, Cairo, EGYPT, World Scientific and Engineering Acad and Soc.

A grid is a collection of distributed computing resources available over a local or wide area network that appear to an end user or application as one large virtual computing system. Grid computing has serious social consequences and is going revolutionize the world of computing. Education is one domain that can be benefited the most from grid computing. The focus on leveraging technologies and products from academic institutions around the world in the areas of e-learning, front-end portals, library automation and digital asset management to promote the sharing and collaboration in the use of technologies, developing content, defining standards and building best-of-breed pedagogical practices is the new upcoming trend. With a lot of work going on in this emerging domain, internationally, India too has joined the bandwagon to contribute to the development in this discipline, This paper aims to present the state-of-the-art of grid computing and attempts to survey India's efforts in developing this upcoming technology. By allowing anyone, anywhere, anytime to easily access supercomputer level processing power and knowledge resources, grids are underpining progress in Indian Education.

Stockinger, H. (2007). "Defining the grid: a snapshot on the current view." Journal of Supercomputing **42**(1): 3-17.

The term "Grid" was introduced in early 1998 with the launch of the book "The Grid. Blueprint for a new computing infrastructure". Since that time many technological

changes have occurred in both hardware and software. One of the most important ones seems to be the wide acceptance of Web services. Although the basic Grid idea has not changed much in the last decade, many people have different ideas about what a Grid really is. In the following article we report on a survey where we invited many people in the field of Grid computing to give us their current understanding.

Stratan, C., A. Iosup, et al. (2008). A performance study of grid workflow engines. Proceedings - IEEE/ACM International Workshop on Grid Computing, Tsukuba.

To benefit from grids, scientists require grid workflow engines that automatically manage the execution of interrelated jobs on the grid infrastructure. So far, the workflows community has focused on scheduling algorithms and on interface tools. Thus, while several grid workflow engines have been deployed, little is known about their performance-related characteristics, and there are no commonly-used testing practices. This situation limits the adoption of the grid workflow engines, and hampers their tuning and their further development. In this work we propose a testing methodology for grid workflow engines that focuses on five characteristics: overhead, raw performance, stability, scalability, and reliability. Using this methodology, we evaluate in a real test environment several middleware stacks that include grid workflow engines, including two based on DAGMan/Condor and on Karajan/Globus. Â© 2008 IEEE.

Tang, F. L., M. L. Li, et al. (2006). "Automatic transaction compensation for reliable grid applications." Journal of Computer Science and Technology **21**(4): 529-536.

As grid technology is expanding from scientific computing to business applications, service oriented grid computing is aimed at providing reliable services for users and hiding complexity of service processes from them. The grid services for coordinating long-lived transactions that occur in business applications play an important role in reliable grid applications. In this paper, the grid transaction service (GridTS) is proposed for dealing with long-lived business transactions. We present a compensation-based long-lived transaction coordination algorithm that enables users to select results from committed sub-transactions. Unlike other long-lived transaction models that require application programmers to develop corresponding compensating transactions, GridTS can automatically generate compensating transactions on execution of a long-lived grid transaction. The simulation result has demonstrated the feasibility of GridTS and effectiveness of the corresponding algorithm. Â© Springer Science + Business Media, Inc. 2006.

Trezentos, P. and A. L. Oliveira (2003). Metrics for grid applicability: A distributed elliptic curve platform assessment. 5th International Conference on Parallel Processing and Applied Mathematics, Czestochowa, POLAND, Springer-Verlag Berlin.

The distribution of computational load among several nodes is an important step for problems requiring High Performance Computing (HPC). This phase is critical because bad decisions could cost time and money. The emergence of heterogeneous networks with resources available for a wide group of users as provided by Grid Computing [3] creates the need for formal processes and metrics to evaluate if the application of the problem to the Grid is feasible. In this article we introduce some auxiliary indicators for measuring the potential applicability of a parallel application to a Grid. Using the measures defined in the internal draft (GWD-I) produced by the Network Measurement WG [6] from the Global Grid Forum and RFC 2330 [7], the authors present some aggregate metrics that are useful in the characterization of the parallel applications that are well adapted to existing grids. The defined auxiliary metrics were useful in the validation of a concrete application that factorizes numbers using the

elliptic curves method (ECM) [5] over a testbed Grid. The results of the application of the metrics to this specific mathematical algorithm are presented.

Truong, H. L., R. Samborski, et al. (2006). Towards a framework for monitoring and analyzing QoS metrics of grid services. e-Science 2006 - Second IEEE International Conference on e-Science and Grid Computing, Amsterdam.

QoS (Quality of Service) parameters play a key role in selecting Grid resources and optimizing resources usage efficiently. Although many works have focused on using QoS metrics, surprisingly few tools support the monitoring and analysis of QoS metrics of Grid services. This paper presents a novel framework which supports the monitoring and analysis of QoS metrics in the Grid. Our approach is that, firstly, we develop a classification of important QoS metrics for Grid services that should be monitored and analyzed. Secondly, sensors are developed to monitor QoS of disparate Grid services by using a peer-to-peer Grid monitoring middleware. The dependencies among Grid services are modeled. Based on that, several techniques are used to analyze QoS metrics of dependent Grid services. © 2006 IEEE.

VÃ¡zquez, C., E. Huedo, et al. (2008). A performance model for federated grid infrastructures. Proceedings of the 16th Euromicro Conference on Parallel, Distributed and Network-Based Processing, PDP 2008, Toulouse.

A performance model, previously proposed to characterize the performance of grid infrastructures, is extended to evaluate federations of grids by aggregating their performance parameters. These parameters can then be used to develop scheduling policies based on them. The new model can be used to take scheduling decisions based on them and hence to aid in the development of scheduling policies. The model has been validated using the performance results obtained in the execution of a high throughput computing application on an enterprise grid composed of Globus Toolkit Web Service resources and a GridGateWay giving access to gLite resources from the EGEE infrastructure. Â© 2008 IEEE.

van Moorsel, A. P. A. (2005). "Grid, management and self-management." Computer Journal **48**(3): 325-332.

Increasingly frequently, Grid software is being proposed as a software platform for operational management of enterprise computing systems, including desktops, data centres and business applications. We argue in this paper that the challenge for future management software is 'enabling automated and business-driven management, for systems with increasing scale, heterogeneity and federation challenges'. Among the currently available middlewares, the Grid platform is undoubtedly best positioned to respond to this challenge, and we review several recent technologies that utilize Grid software for management purposes. However, Grid software also has its limitations. Its software architecture provides no particular support to building self-managing distributed systems and is therefore most viable for traditional hierarchical management systems, providing rudimentary self-management capabilities through control loops. Â© The Author 2005. Published by Oxford University Press on behalf of The British Computer Society. All rights reserved.

Vanderster, D. C., N. J. Dimopoulos, et al. (2009). "Resource allocation on computational grids using a utility model and the knapsack problem." Future Generation Computer Systems **25**(1): 35-50.

This work introduces a utility model (UM) for resource allocation on computational grids and formulates the allocation problem as a variant of the 0-1 multichoice multidimensional knapsack problem. The notion of task-option utility is introduced,

and it is used to effect allocation policies. We present a variety of allocation policies, which are expressed as functions of metrics that are both intrinsic and external to the task and resources. An external user-defined credit-value metric is shown to allow users to intervene in the allocation of urgent or low priority tasks. The strategies are evaluated in simulation against random workloads as well as those drawn from real systems. We measure the sensitivity of the UM-derived schedules to variations in the allocation policies and their corresponding utility functions. The UM allocation strategy is shown to optimally allocate resources congruent with the chosen policies. Â© 2008 Elsevier B.V. All rights reserved.

Vaquero, L. M., L. Rodero-Merino, et al. (2009). "A Break in the Clouds: Towards a Cloud Definition." Computer Communication Review **39**(1): 50-55.

This paper discusses the concept of Cloud Computing to achieve a complete definition of what a Cloud is, using the main characteristics typically associated with this paradigm in the literature. More than 20 definitions have been studied allowing for the extraction of a consensus definition as well as a minimum definition containing the essential characteristics. This paper pays much attention to the Grid paradigm, as it is often confused with Cloud technologies. We also describe the relationships and distinctions between the Grid and Cloud approaches.

Vazquez, T., E. Huedo, et al. (2007). Evaluation of a utility computing model based on the federation of grid infrastructures. 13th International Euro-Par Conference on Parallel Processing, Rennes, FRANCE, Springer-Verlag Berlin.

Utility computing is a service provisioning model which will provide adaptive, flexible and simple access to computing resources, enabling a pay-per-use model for computing similar to traditional utilities such as water, gas or electricity. On the other hand, grid technology provides standard functionality for flexible integration of diverse distributed resources. This paper describes and evaluates an innovative solution for utility computing, based on grid federation, which can be easily deployed on any infrastructure based on the Globus Toolkit. This solution exhibits many advantages in terms of security, scalability and site autonomy, and achieves good performance, as shown by results, mainly with computeintensive applications.

Venugopal, S., R. Buyya, et al. (2003). A Grid service broker for scheduling e-Science applications on global data Grids. 1st International Workshop on Middleware for Grid Computing held in Conjunction with the ACM/IFIP/USINIX Middleware Conference, Rio de Janeiro, BRAZIL, John Wiley & Sons Ltd.

The next generation of scientific experiments and studies, popularly called e-Science, is carried out by large collaborations of researchers distributed around the world engaged in the analysis of huge collections of data generated by scientific instruments. Grid computing has emerged as an enabler for e-Science as it permits the creation of virtual organizations that bring together communities with common objectives. Within a community, data collections are stored or replicated on distributed resources to enhance storage capability or the efficiency of access. In such an environment, scientists need to have the ability to carry out their studies by transparently accessing distributed data and computational resources. In this paper, we propose and develop a Grid broker that mediates access to distributed resources by: (a) discovering suitable data and computational resources sources for a given analysis scenario; (b) optimally mapping analysis jobs to resources; (c) deploying and monitoring job execution on selected resources; (d) accessing data from local or remote data sources during job execution; and (e) collating and presenting results. The broker supports a declarative and dynamic parametric programming model for creating Grid applications. We have

used this model in Grid-enabling a high-energy physics analysis application (the Belle Analysis Software Framework). The broker has been used in deploying Belle experimental data analysis jobs on a Grid testbed, called the Belle Analysis Data Grid, having resources distributed across Australia interconnected through GrangeNet. Copyright (c) 2005 John Wiley & Sons, Ltd.

Vykoukal, J., M. Wolf, et al. (2009). "Services Grids in industry on-demand provisioning and allocation of Grid-based business services." Service-Grids in der Industrie - On-Demand-Bereitstellung und Nutzung von Grid-Basierten Services in Unternehmen **51**(2): 206-214.

Over the last few years Grid computing has attracted considerable attention from the industry, because it offers opportunities for new on-demand business services for enterprises. In this article, we depict the current trend of enterprises to source Grid services offered by third-party utility providers on a use-on-demand, pay-per-use basis, thus leading to faster IT response to changing business needs. We outline the emerging expansion of Grid computing In industry and discuss the significant challenges still needing to be solved to further promote the adoption of Grid technology in the business domain.

Weiss, M. B. H. (2004). Models of grids: Cost and standards implications. GECON 2004 - 2004 1st IEEE International Workshop on Grid Economics and Business Models, Seoul.

The emergence of grid computing as a significant new model of computing is discussed. The centrally formed grids provides the platform for commercial utility computing system. The development of an appropriate middleware leads to a hybrid-captive grids in which centrally owned machines are dedicated for personal use. The results show that the centrally owned machines can be used in support of commercial utility service.

Wu, J., L. Deng, et al. (2008). "Pricing strategies for grid computational economy." Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University **26**(4): 497-502.

Purpose. There exist already many pricing strategies for grid computational economy. The full paper begins with the discussion of a number of strategies mentioned in Refs.2 through 7. We present our strategy which we hope is better than those discussed. In the full paper, we explain our strategy in some detail; here we just give a briefing. Our strategy consists of price negotiation and price regulation. The explanation in the full paper centers around two topics: (1) TÃ tonnement process, which implements the price negotiation between GSC (Grid Services Consumer) and GSP (Grid Services Provider); in the TÃ tonnement process, pricing is dependent on the demand for resources in the market; (2) price regulation, which is based on the Ramsey Pricing and the Price cap proposed by Stephen Littlechild. Finally we performed computer simulations to compare the performance of our strategy with that of the auction model in Ref.10 by Rich Wolski et al. The computer simulation results, shown in Table 1 in the full paper, show preliminarily that our strategy is better than that of Ref.10 in that: (1) the throughput is higher by 28.6%, and (2) the average utilization ratio of CPU increases by 18.08%.

Wu, S., W. Zhu, et al. (2008). VMGrid: A virtual machine supported grid middleware. 2008 IEEE International Conference on Granular Computing, GRC 2008, Hangzhou.

Many grid middleware such as Globus Toolkit and GPE have started to support virtual machines. They implement the remote control of virtual machine by grid service and make use of virtual machines' function of quick deployment to construct the execution environment of grid application fast. In order to support virtual machine in grid environment, VMGrid (Virtual Machine Grid) is presented. All services in VMGrid are

implemented based on secure shell host protocol (SSH) while Globus Toolkit's Virtual Workspace is on WSRF. This paper gives a clear description of VMGrid. Some comparisons with Globus Toolkit's Virtual Workspace in performance are made in this paper. Performance results show that VMGrid virtual workspace atomic services are more effective than that of Globus.

Xie, X., J. Huang, et al. (2008). Pricing strategies in grid market: Simulation and analysis. Proceedings - 7th International Conference on Grid and Cooperative Computing, GCC 2008, Shenzhen.

The most important concept of a market is price - the terms on which the trading objects (products or services) are exchanged. In an ideal market, the price is a reflection of the current state of the market, and therefore it should be dynamic. From the seller perspective, using an effective dynamic pricing strategy can make a difference between making a profit or loss in a makret. In a grid marketplace, the strategy is even more important given it is very dynamic and volatile nature. In this paper, we describe and review three different pricing strategies for the grid market: Demand-Response, Revenue-Response and Price-Maximizing strategy, and analyze their behavior in different market conditions using simulation approach. Our results show that two of the strategies are able to generate high revenue in various market conditions Â© 2008 IEEE.

Yeo, C. S. and R. Buyy (2006). "A taxonomy of market-based resource management systems for utility-driven cluster computing." Software - Practice and Experience **36**(13): 1381-1419.

In utility-driven cluster computing, cluster Resource Management Systems (RMSs) need to know the specific needs of different users in order to allocate resources according to their needs. This in turn is vital to achieve service-oriented Grid computing that harnesses resources distributed worldwide based on users' objectives. Recently, numerous market-based RMSs have been proposed to make use of real-world market concepts and behavior to assign resources to users for various computing platforms. The aim of this paper is to develop a taxonomy that characterizes and classifies how market-based RMSs can support utility-driven cluster computing in practice. The taxonomy is then mapped to existing market-based RMSs designed for both cluster and other computing platforms to survey current research developments and identify outstanding issues. Copyright Â© 2006 John Wiley & Sons, Ltd.

Yongwei, W., Y. Guangwen, et al. (2003). Grid computing pool and its framework. Parallel Processing Workshops, 2003. Proceedings. 2003 International Conference on.

The aim of grid computing pool (GCP) is to integrate heterogeneous large-scale computing devices and famous computing tools into a virtual supercomputer by using grid technology. It makes supercomputing services Internet accessible easily and widely. Performance analysis shows that GCP improves service quality and utilization of these computing resources greatly. By implementing grid computing engine interface (GCEI), it interfaces with multiple heterogeneous mathematical computing tools over the grid. The overall GCP description and its framework are presented in detail. Related work and implementation experience are also reported.

Youseff, L., M. Butrico, et al. (2008). Toward a unified ontology of cloud computing. Grid Computing Environments Workshop, GCE 2008, Austin, TX.

Progress of research efforts in a novel technology is contingent on having a rigorous organization of its knowledge domain and a comprehensive understanding of all the relevant components of this technology and their relationships. Cloud Computing is

one contemporary technology in which the research community has recently embarked. Manifesting itself as the descendant of several other computing research areas such as Service-Oriented Architecture, distributed and grid computing, and virtualization, cloud computing inherits their advancements and limitations. Towards the end-goal of a thorough comprehension of the field of cloud computing, and a more rapid adoption from the scientific community, we propose in this paper an ontology of this area which demonstrates a dissection of the cloud into five main layers, and illustrates their interrelations as well as their inter-dependency on preceding technologies. The contribution of this paper lies in being one of the first attempts to establish a detailed ontology of the cloud. Better comprehension of the technology would enable the community to design more efficient portals and gateways for the cloud, and facilitate the adoption of this novel computing approach in scientific environments. In turn, this will assist the scientific community to expedite its contributions and insights into this evolving computing field.

Yu, J.-j., Y.-z. Zheng, et al. (2007). "Resource scheduling algorithm for computing grid based on utility optimization." Journal of Computer Applications **27**(3): 541-9.

Concerning the time and budget restriction in resource scheduling, a utility function model was proposed by analyzing the relationship between the time and budget of task and the user utility function. Then, based on the utility function model, a utility optimization based heuristic resource scheduling algorithm was put forward for task farming application model. Simulation results show that our proposed algorithm provides user with more utilities than cost-time optimization algorithm

Yu, J. and R. Buyya (2005). "A taxonomy of workflow management systems for Grid computing." Journal of Grid Computing **3**(3-4): 171-200.

With the advent of Grid and application technologies, scientists and engineers are building more and more complex applications to manage and process large data sets, and execute scientific experiments on distributed resources. Such application scenarios require means for composing and executing complex workflows. Therefore, many efforts have been made towards the development of workflow management systems for Grid computing. In this paper, we propose a taxonomy that characterizes and classifies various approaches for building and executing workflows on Grids. We also survey several representative Grid workflow systems developed by various projects world-wide to demonstrate the comprehensiveness of the taxonomy. The taxonomy not only highlights the design and engineering similarities and differences of state-of-the-art in Grid workflow systems, but also identifies the areas that need further research. Â© Springer Science+Business Media, Inc. 2006.

Yu, J., R. Buyya, et al. (2008). Workflow scheduling algorithms for grid computing. Studies in Computational Intelligence. F. X. Xhafa and A. A. Abraham. **146:** 173-214.

Workflow scheduling is one of the key issues in the management of workflow execution. Scheduling is a process that maps and manages execution of inter-dependent tasks on distributed resources. It introduces allocating suitable resources to workflow tasks so that the execution can be completed to satisfy objective functions specified by users. Proper scheduling can have significant impact on the performance of the system. In this chapter, we investigate existing workflow scheduling algorithms developed and deployed by various Grid projects. Â© 2008 Springer-Verlag Berlin Heidelberg.

Yu, J., R. Buyya, et al. (2005). Cost-based scheduling of scientific workflow applications on utility grids. Proceedings - First International Conference on e-Science and Grid Computing, e-Science 2005, Melbourne.

over the last few years, Grid technologies have progressed towards a service-oriented paradigm that enables a new way of service provisioning based on utility computing models. Users consume these services based on their QoS (Quality of Service) requirements. In such "pay-per-use" Grids, workflow execution cost must be considered during scheduling based on users' QoS constraints. In this paper, we propose a cost-based workflow scheduling algorithm that minimizes execution cost while meeting the deadline for delivering results. It can also adapt to the delays of service executions by rescheduling unexecuted tasks. We also attempt to optimally solve the task scheduling problem in branches with several sequential tasks by modeling the branch as a Markov Decision Process and using the value iteration method. Â© 2005 IEEE.

Yuan, L., G. Zeng, et al. (2006). A Grid resource price-adjusting strategy based on price influence model. Proceedings - Fifth International Conference on Grid and Cooperative Computing, GCC 2006, Hunan.

In the market-oriented Grid system, the price management of Grid resources is a challenging undertaking as most resources are interdependent. So we discuss the dependent relations among Grid resources, and build a "Price Influence Model" which reflects the price dependence of resources. Furthermore, a price-adjusting strategy based on the "Price Influence Model" is proposed. The strategy focuses on adjusting the unreasonable prices of partial resources, so as to achieve equilibrium prices of all resources. Finally, we illustrate the utility of our work by one example. Â© 2006 IEEE.

Zanikolas, S. and R. Sakellariou (2004). Towards a monitoring framework for worldwide Grid information services. 10th International Euro-Par Conference on Parallel Processing, Pisa, ITALY, Springer-Verlag Berlin.

Despite important advances in Grid computing, scalability and interoperability issues obstruct the integration of the existing isolated grids into the Grid. We propose a worldwide monitoring framework which forms the basis for building worldwide grid information services that can have a global notion of the Grid. This work motivates the monitoring framework in terms of applications and challenges, sets its architecture and design space, and offers indicative performance results in a variety of hypothetical Grid instances.

Zhang, L. J., H. Li, et al. (2004). "Toward a business process grid for utility computing." IT Professional **6**(5).

The incorporation of underused computing capacity into existing grid computing technologies to solve business problems and provide IT-level infrastructure to support business is discussed. For the grid, it is recommended to develop a flow technology capable of adapting to the dynamic grid environment and changing requirements of grid applications. Due to the integration of grid computing and Web services, much of the work in Web services flow greatly influences grid service flow and the business grid. A common approach to address the interoperability issue between two state management frameworks, one in grid services and another in business processes is advocated for further benefits.

Zhang, W. Z., X. R. Liu, et al. (2006). "Trust-driven job scheduling heuristics for computing grid." Tongxin Xuebao/Journal on Communication **27**(2): 73-79.

Currently existing grid scheduling algorithms largely ignore the impact of trust mechanism, which made their applicability in a realistic environment rather doubtful. Based on the grid trust model and trust utilization functions, a trust-driven grid job scheduling problem was proposed. Trust-driven Min-min and Sufferage job scheduling heuristics were presented, which were extended based on traditional job scheduling strategies. The algorithms were evaluated with large scale simulation. Simulation result demonstrates trust-driven heuristics perform better than time-driven heuristic on average trust utility and total trust utility.

Zhang, Y., L. Lin, et al. (2006). "Resource allocation mechanism providing trust and incentive in grid." Ruan Jian Xue Bao/Journal of Software **17**(11): 2245-2254.

In grid environment, when every participating node tries to maximize its own utility, the overall utility of the collaboration drops. In the worst case scenario, grid resources are easily depleted due to selfish users taking free rides without offering any sharing resource. Certain resource management scheme has to be implemented on grid systems to ensure them working properly and to achieve better scalability. In this paper, the relationship of resources incentive and allocation in the grid is discussed. In grid computing, the evaluation of shared resource is the key to allocate and manage them. To address the issue, the price strategy of sellers and an adaptive trust-incentive compatible resource allocation mechanism ATIM (adaptive trust-incentive compatible resource allocation) are presented to ensure the balance of supply and demand, promote the users to share valuable resource, and maximize resource utility. Intending for a secure collaborative environment, the framework considers the factors of trust and economy to ensure the steady development of grid resource market. This approach is also evaluated by comprehensive experiments in CROWN (China research and development environment over wider-area network).