A proposed Semantic Description Framework for Semantic Grid Resources

Tamer A. Farrag^{1,2}, Mohamed T. Faheem^{1,3}

¹ Department of Electrical Engineering, Faculty of Engineering, Taif University, Saudi Arabia

² Department of Computers and Systems Engineering , Faculty of Engineering, Mansoura University, Egypt

³ Department of Computers and Automatic Control Engineering, Faculty of Engineering, Tanta University, Egypt

ABSTRACT— The grid is the base of the modern trends in the field of computing systems. The grid systems suffer from the problem of weakness of the current data models which are used to define the gird resources. These models suffer from the lack of semantics which leads to resources wasting. The lack of semantics yields in misunderstanding of resources information and makes the process of inferring new relations among the grid resources almost not available. The use of the semantic data model makes the information about grid resources not only machine-readable (like the current grid systems) but also machine-understandable. The Semantic Grid refers to an approach used to grid computing in which information, computing resources and services are described using the semantic data model. The aim of this work is to use the semantic web technologies such as RDF , RDFS, OWL to obtain a new framework to redefine the different grid resources in a semantic manner.

Index Terms— Grid , Semantic , resource description ,RDF , web 3.0 , machine-understandable

_ _ _ _ _ _ _ _ _ _ _ _

1. INTRODUCTION

The grid can be considered as the base of the I modern trends in the field of computing systems. The grid systems suffer from the problem of weakness of the current data models which are used to define the gird resources. These models suffer from the lack of semantics which leads to resources wasting. The lack of semantics yields in misunderstanding of resources information and makes the process of inferring new relations among the grid resources almost not available. The use of the semantic data model makes the information about grid resources not only machine-readable (like the current grid systems) but also machine-understandable. Therefore, the process of importing new technologies of web 3.0 and the semantic web produces a new enhanced generation of the grid called the semantic grid. The Semantic Grid refers to an approach used to grid computing in which information, computing resources and services are described using the semantic data model. The proposed semantic model will be used to compute-intensive, write-intensive and read-intensive applications.

The rest of this paper is organized as follows: In section 2, a brief background about the developing of the grid computing is introduced. In section 3, we focus on the new trend of the grid computing which is called Semantic Grid. In section 4, the current Grid Resources Description Framework is surveyed. In section 5, we introduce our vision about a proposed new Semantic grid resources description framework.. In section 6, we introduce our prepared testing bed. Finally , we conclude this paper.

2. GRID COMPUTING

Due to the dire need to get over the problem of centralization and the high cost of supercomputers, the Grid Computing appeared in 1990s[1]. The Grid consists of a large number of PCs distributed geographically and communicated together through a special network. The Grid is the infrastructure of the distributed processing model in which the unused processing power of PCs is imposed by the other PCs, which have large processing needs. The research in the field of the grid computing results many new concepts that enrich the computer science such as Internet computing , mobile computing [2] and recently cloud computing [3][4]. All of these concepts is derived from the idea of grid computing.

In addition, the unexpected ongoing of the Internet and the huge number of PCs and resources attached to the Internet rises a question "Why we do not apply this huge Computing Power on a large scale?" Although that more 10 billion computers worldwide are connected to the Internet (according to IMS Research.[5]), researches assert that about 90 percent of the time these computers are on, they are idle—that is, not actively performing computing tasks. The sum of these unused computational powers is 25000 times more powerful than the most powerful supercomputer all over the world.

So, the research in the field of grid computing become affected by the development of the Internet researches. For example, All the modern approaches that deals with the problem of enabling resource sharing of geographically diverse computational resources, use the Web Services. The new trend in the Internet researches is a about the next generation of the web which called web 3.0 or the semantic web. Web 3.0 is a set of different technologies that can enormously change the existed form and technique of the current web functions. With these new technologies, the web would become much smarter and easier for people to use. There is no standard definition of web 3.0 technology; however, this concept is defined by different experts in various ways.

3. SEMANTIC GRID

In 2001, Berners-Lee said, "The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation"[6]. In other words, semantic web is an evolving extension of the current www in which the semantics of information and services on the web is defined, making it possible for the web to understand and satisfy the requests of people and machines to efficiently use the web content [7][8]. So, semantic web aims to convert the web contents from machine readable to machine understandable [10][11][12].

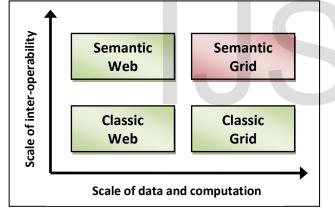


Figure 1: From the current Grid to the semantic grid

As shown in Figure 1, the process of importing the new technologies [13][14][15] of the web 3.0 and the semantic web produces a new enhanced generation of the grid called semantic grid[17]. The Semantic Grid refers to an approach to grid computing in which information, computing resources and services are described using the semantic data model [18][19]. As shown Figure 1, Semantic Grid combines higher inter-operability (of sematic web) with greater computational facilities (of the grid).

The main contributions of using semantic web technologies in the grid context are : (1) redefine the grid resources and services in a semantic manner (2) replace the use of the conventional web services by the new concept of semantic web services. So, there is a need to propose a new framework or upper ontology[20][21] to describe the grid

resources. This proposal benefit from the semantic web languages and the modern researches in field of semantic web services[22][23][24]. Figure 2 summarizes the main research areas in the computing systems. This research can be considered as a contribution is fields of data management and resource discovery.

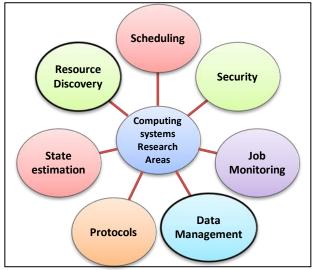


Figure 2: Computing systems Research Areas taxonomy

4. GRID RESOURCES DESCRIPTION FRAMEWORK

The grid resources description frameworks (GRDFs) can be categorized into two main categories according to the used technologies. They can categorized to semantic and non-semantic frameworks. In non-semantic ones, the resource provider and the resource customers descript the available or required resource in a nonstandard way. For example, the resource provider can say that "I have 32MB free Memory" and on the other hand the resource provider can say that " I need 16MB RAM". In this example the resource matchmaking system cannot infers that the mean of the "free memory" equivalent to "free Ram" except if the system administrator provide this information manually.

The main idea of the semantic web is to extend the current human-readable web by encoding some of the semantics of resources in a machine-processable form [25]. Moving beyond syntax opens the door to more advanced applications and functionalities on the Web. So, the semantic GRDFs will enhance the way of grid resources description which enable the resources management and matchmaking systems to infer new non-provided relations, unify the ways of resource advertising and resource requesting. The non-semantic GRDFs use Attribute based resource description model and the semantic GRDFs use ontology –based resource description model.

Next in this section, we will survey the most important non-semantic resource management systems:

1. Condor: Cycle stealing technology for high throughput computing: This project is labeled to be one of the most popular systems. The Condor is a High Throughput Computing System, which has scavenged otherwise wasted CPU cycles from desktop workstations for more than a decade. These inconsistently available resources have proven to be a significant source of computational power, enabling scientists to solve ever more complex problems. Condor efficiently harnesses existing resources, reducing or eliminating the need to purchase expensive supercomputer time or equivalent hardware [26]. The requests and offers of the resource are described in the Condor classified advertisement (ClassAd) language [27]. Class Ads use a semi-structured data model for resource description. The Class Ad language includes a query language as part of the data model, allowing advertising agents to specify their compatibility by including constraints in the offers and requests of their resources. Moreover, Condor can be considered as computational Grid with a flat organization. The resource discovery includes centralized queries with periodic push dissemination and therefore the scheduler is centralized.

2. European DataGrid: The European DataGrid Project focuses on the development of middleware services in order to enable distributed analysis of physics data. The core middleware system is the Globus toolkit with hooks for data Grids. Data on the order of several petabytes will be distributed in a hierarchical fashion to multiple sites worldwide. Accordingly, global namespaces are required to handle the creation of and the access to distributed and replicated data items. Moreover, special workload distribution facilities will balance the analysis jobs from several hundred physicists to different places in the Grid in order to have maximum throughput for a large user community. Application monitoring as well as collecting of user access patterns will provide information for access and data distribution optimization.

3. The DataGrid project has a multi-tier hierarchical RMS organization. For example, tier-0 is CERN, which stores almost all relevant data; several tier-1 regional centers (in Italy, France, U.K., U.S.A., Japan) will support smaller amounts of data, and so on. It has an extensible schema-based resource model with a hierarchical namespace organization. It does not offer any QoS and the resource information store is expected to be based on an LDAP network directory. Resource dissemination is batched and periodically pushed to other parts of the Grid.

4. Globus: A toolkit for Grid computing :The open source Globus Toolkit is a fundamental enabling technology for the "Grid," helping people to share computing power, databases, and other tools securely online across corporate, institutional, and geographic boundaries without sacrificing local autonomy. The toolkit includes software services and libraries for resource monitoring, discovery, and management, plus security and file management. In addition to being a central part of science and engineering projects that total nearly a halfbillion dollars internationally, the Globus Toolkit is a substrate which leads IT companies build significant Grid commercial products. The toolkit includes software for security, information infrastructure, resource management, data management, communication, fault detection, and portability. It is packaged as a set of components that can be used either independently or collectively to develop applications. Every organization has its unique modes of operation, and collaboration among multiple organizations which can be hindered by incompatibility of resources such as data archives, computers, and networks. The Globus Toolkit is conceived to remove obstacles that prevent seamless collaboration. Its core services, interfaces and protocols allow users to access remote resources as if they were located within their own machine room that simultaneously preserving local control over the one who can use resources and when. The Globus Toolkit has grown through an open-source strategy similar to the Linux operating systems, and distinct from proprietary attempts at resource-sharing software. This encourages broader, more rapid adoption and leads to greater technical innovation, as the open-source community provides continual enhancements to the product.

5. Nimrod/G: Resource broker and economy grid: Nimrod is a tool that manages the execution of parametric studies across distributed computers. It takes responsibility for the overall management of an experiment, as well as the low-level issues of distributing files to remote systems, above all performing the remote computation and gathering the results. The system then uses this information to transport the necessary files and schedule the work on the first available machine. Nimrod/G is a Grid aware version of Nimrod. It takes advantage of features supported in the Globus toolkit such as automatic discovery of allowed resources. Nimrod/G is being used as a scheduling component in a new framework called Grid Architecture for Computational Economy (GRACE) which is based on using economic theories for a Grid resource management system.

5. PROPOSED SEMANTIC GRID RESOURCES DESCRIPTION FRAMEWORKS

In order to design a semantic grid, the first step is to design an upper ontology to define any grid aspect.

Therefore, a survey about the grid components and the grid workflow should be prepared to count the main concepts and their properties that should be included in this upper ontology. Simply , the grid consists of a number of resources that cooperate to perform the different jobs requested by the grid users. So, The three main concepts in our proposed grid upper ontology are Grid Resource, Grid User and Grid Job as shown in Figure 3 .

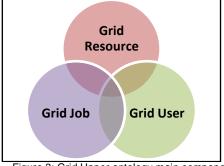


Figure 3: Grid Upper ontology main component

5.1. Semantic Grid Resources Definition

In this paper , we focus on the semantic definition of the grid resources A Grid resource is a Grid entity that provides some capabilities to a consumer. different resources could provide similar capabilities but with different quality of service aspects. The resource capabilities are required to be presented in such a way that a consumer can easily discover a resource or a resource ensemble with needed capabilities. By employing this, we attain the following:

- Grid resource management systems make a large step towards compatibility with the Semantic Web and Grid resources descriptions become web-understandable.
- The resource provider can have maximum freedom to describe resources with different levels of complexity and completeness.
- XML-schema datatypes can be exploited for resource description.
- Complex resource matching is possible based on subsumption relationship.
- A conceptual defnition of resources in a more natural way is possible based on the restriction over the resource attributes, and a semantics level of agreement between resource provider and consumer can be achieved.
- After having a conceptual and flexible resource description, a resource broker can categorize the resource ensembles and devise alternative options.
- Most promisingly, clients can express complex requests in a simple human as well as machine understandable format. Also the system can fulfill more resource requests than without ontology language. For example,

in a system without ontology language, a request for *Unix* system fails if the term *Unix* is not specified. Using an ontology this might however be successful.

• Spelling or typing errors in descriptions and requests are prevented by using a controlled vocabulary.

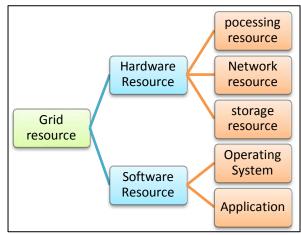


Figure 4: Grid Resource main sub concepts.

Figure 4 shows uncompleted taxonomy of the grid resources sub concepts which should be included in any upper ontology to define and descript and resource in the semantic grid. We select OWL [28] for our current work. OWL is a good language for providing more complex constraints on the types of resources and their properties. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDFS. OWL has three increasingly expressive sublanguages: OWL Lite, OWL DL, and OWL Full.

Annotation Properties Individuals OVILVE DL Duery Ontionst SPAROL Duery Ontobogy Differences Cases Network United	💠 🔄 🐨 GridResource 🔹 Exactly for endly						
Costs herechy flobing Clip is Image: Strange Resource Image: Strange Resource Pocessing Resource Strange Resource Strange Resource Strange Resource Application Operating System Contract Resource Image: Strange Resource Strange Resource Strange Resource Application Operating System Contract Resource Strange Resource<th></th><th></th><th></th>							
Thing GridResource HardwareResource HoressingResource StorageResource Application OperatingSystem	Class hierarchy: Nothing DEIRS	A STATE OF	080				
	HardwareResource NetworkResource PocessingResource StorageResource SoftwareResource Application	Equation to 💭 Balactions of 🕐	a:H				

Figure 5: Protégé snapshot for grid resource ontology creation

We use Protégé [29] to write our grid upper ontology as shown in Figure 5. Protégé is a free, open-source platform that provides a growing user community with a suite of tools to construct domain models and knowledgebased applications with ontologies. Protégé enables users to build ontologies for the Semantic Web, in particular in the W3C's OWL. An OWL ontology may include descriptions of classes, properties and their instances.

Given such an ontology, the OWL formal semantics specifies how to derive its logical consequences, i.e. facts not literally present in the ontology, but entailed by the semantics. These entailments may be based on a single document or multiple distributed documents that have been combined using defined OWL mechanisms. Figure 6 shows a part of the OWL file that defines the contents of the Grid resource ontology.



</Ontology>

Figure 6: part of OWL file of Grid Resource Ontology

6. TESTING BED

In order to test our proposals, we construct a small Grid consist of 6 PCs. Alchemi is used to implement this testing grid. However, there is a distinct lack of serviceoriented architecture-based grid computing software in this space. To overcome this limitation, a Windows-based desktop grid computing framework called Alchemi is developed and implemented on the Microsoft .NET Platform. The Microsoft .NET Framework is the state of the art development platform for Windows and offers a number of features which can be leveraged for enabling a computational desktop grid environment on Windowsclass machines. Next, we present the main features of Alchemi and the main terminologies that are used in it.

6.1. ALCHEMI

Alchemi is conceived with the aim of making grid construction and development of grid software as easy as possible without sacrificing flexibility, scalability, reliability and extensibility. The key features supported by Alchemi are:

- Internet-based clustering [30][31] of Windows-based desktop computers.
- Dedicated or non-dedicated (voluntary) execution by individual nodes.
- Object-oriented grid application programming model (fine-grained abstraction).
- file-based grid job model (coarse-grained abstraction) for grid-enabling legacy applications.
- web services interface supporting the job model for interoperability with custom grid middleware e.g. for creating a global, cross-platform grid environment via a custom resource broker component.

According to Alchemi, Grids are constructed using three types of distributed components (or nodes). They are named according to their roles with respect to a grid application .:

1. Manager :The Manager manages the execution of grid applications and provides services associated with managing thread execution. It is deployed as an executable. An optional sub-component of the Manager is the Cross Platform Manager, which is deployed as a web service.

2. Executor: The Executor executes individual grid threads and provides services associated with executing threads. It is deployed as an executable. An Executor can be configured to be dedicated (meaning the Manager initiates thread execution directly) or non-dedicated (meaning that thread execution is initiated by the Executor on a volunteer basis via a screen saver or some other user-defined options.)

3. **Owner** : The Owner owns an application and provides services associated with the ownership of an application (and its constituent threads). The Owner is implicitly created by the Alchemi API.

Figure 7 illustrate the relation between the different types of Alchemi components.

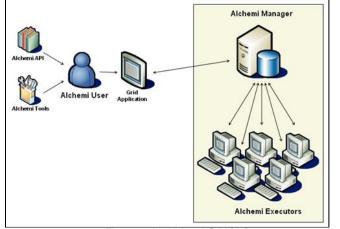


Figure 7: An Alchemi Grid [32]

6.2. TESTING GRID CONFIGURATION

As mention before, our testing grid consist of 6 PCs. We configure one PC as a grid manager and the rest PCs as executor as shown in Figure 8.

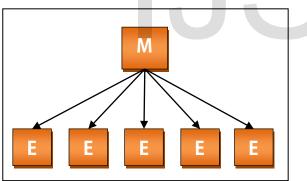


Figure 8: Testing Grid Configuration

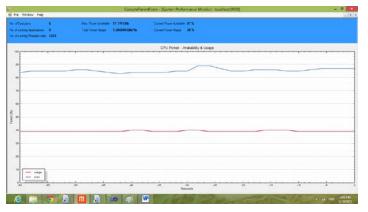


Figure 9: Alchemi Console (System Performance Monitor Window)

We use the Alchemi console application to monitor the process of grid establishing as shown in Figure 9. We use some of applications examples provided by Alchemi developer to ensure the work of the testing grid as shown in Figure 10.

ConsoleF	ParentForm	- [Alchemi (Console]			×
💀 File Action Tools View Window Help)				-	. 8 ×
🔜 🖸 🖄 🗙 📾 🔛						
⊡- 11 Console Root ⊡- 50 Users and Groups	44a	4h.	44a	49a	44h	^
Users Groups	0	1	2	3	4	
Permissions Executors Applications Pl Calculator - Alchemi sample Pl Calculator - Alchemi sample Pl Calculator - Alchemi sample	*	1		*		
	5	6	7	8	9	
	٠	\$	٠	\$		
	10	11	12	13	14	
	*	*	*	100	*	
	15	16	17	18	19	
	₩.	*	*	*	*	
	20	21	22	23	24	
Connected to grid as admin@localbost:9000	4 44.	4 24	d h.	dh .	@ \$_	~

Figure 10: Alchemi Console (Applications tab)

7. CONCLUSION

In this paper, the concept of the semantic grid is introduced. The main target of this paper is to introduce the first step to build a new framework to describe the grid resources in a semantic manner. It is represented in preparing a new upper ontology for the semantic grid resources. Also, we introduce our testing grid which will be used to apply this new framework.

8. REFERENCES

- [1] I.Foster and C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure: Morgan Kaufmann", 1999.
- [2] Pieter Simoens, Filip De Turck, Bart Dhoedt, and Piet Demeester," Remote Display Solutions for Mobile Cloud Computing", IEEE computer society, vol 44 Issue 8, pp.46-53,2011
- [3] Moreno-Vozmediano , "Key Challenges in Cloud Computing to Enable the Future Internet of Services", accepted for publication in IEEE Internet Computing ,2012
- [4] Nathaniel Borenstein and James Blake ,"Cloud Computing Standards: Where's the Beef?", IEEE Internet Computing, vol 15 issue 3, pp74-78, 2011.
- [5] http://imsresearch.com/pressrelease/Internet_Connected_Devices_Approaching_10_Billi on_to_exceed_28_Billion_by_2020&cat_id=190&type=Lates tResearch
- [6] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web", Scientific American Magazine, vol. 284, no. 5, pp. 34–43, 2001.

2517

- [7] Michael C. Daconta, Leo J. Obrst and Kevin T. Smith, "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management", Book, Wiley Publishing 2003.
- [8] G. Antoniou and F. van Harmelen, "A Semantic Web Primer", Book (2nd edition), The MIT Press 2008.
- [9] H. Stuckenschmidt and F. van Harmelen, "Information Sharing on the Semantic Web", Book, Springer 2004.
- [10] L. Li and I. Horrock. "A software framework for matchmaking based on semantic web technology". In Proceeding of 12th International World Wide Web Conference on E-Services and the Semantic Web (ESSW 2003), 2003.
- [11] Jorge Cardoso, "Semantic Web Services: Theory, Tools, and Applications", Book, IGI Global 2007.
- [12] M. Burstein, C. Bussler, M. Zaremba, T. Finin, M. N. Huhns, M. Paolucci, A. P. Sheth, and S. Williams, "A Semantic Web Services Architecture", IEEE Internet Computing, vol. 9, no. 5, pp. 72–81, 2005.
- [13] Resource Description Framework (RDF), W3C Recommendation 2004, online :http://www.w3.org/RDF/
- [14] RDF Vocabulary Description Language 1.0: RDF Schema (RDFS), W3C Recommendation 2004, online :http://www.w3.org/TR/rdf-schema/
- [15] Web Ontology Language (OWL), W3C Recommendation 2004, online: http://www.w3.org/2004/OWL/, 2004.
- [16] D. De Roure, Y. Gil, and J. A. Hendler, "Guest editors' introduction: E-Science," IEEE Intelligent Systems, vol. 19, pp. 24 - 25 2004.
- [17] C. Wroe, C. A. Goble, M. Greenwood, P. Lord, S. Miles, J. Papay, T. Payne, and L. Moreau, "Automating Experiments Using Semantic Data on a Bioinformatics Grid," IEEE Intelligent Systems, vol. 19, pp. 48-55, 2004.
- [18] S. Cox, L. Chen, S. Campobasso, M. Duta, M. Eres, M. Giles, C. Goble, Z. Jiao, A. Keane, G. Pound, A. Roberts, N. Shadbolt, F. Tao, J. Wason, and F. Xu, "Grid Enabled Optimisation and Design Search (GEODISE)." UK e-Science All Hands Meeting 2002 Sheffield, UK, 2002.
- [19] L. Pouchard, S. Bechhofer, B. Matthews, J. Myers, D. Snelling, and Y. Sure, "A Virtual Organisation Ontology", 2006, http://www.csm.ornl.gov/~7lp/onto-library/DagstuhlVirtualOrganization1.0.owl, (2006).
- [20] Dieter Fensel, " Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce", Springer-Verlag, 2001.
- [21] John Davies, Dieter Fensel and Frank van Harmelen ,"TOWARDS THE SEMANTIC WEB: Ontology-driven Knowledge Management", Book, Wiley Publishing 2003.
- [22] Ed Ort, "Service-Oriented Architecture and Web Services: Concepts, Technologies, and Tools", Sun Microsystems Technical Article, April 2005, online: http://java.sun.com/developer/technicalArticles/WebServic es/soa2/.
- [23] Andreas Heß, Eddie Johnston and Nicholas Kushmerick ,"ASSAM: A Tool for Semi-Automatically Annotating Semantic Web Services", in Proceedings of 12th

International Conference on Web Technologies, pp. 470–475,2008

- [24] Abhijit A. Patil, Swapna A. Oundhakar, Amit P. Shethand Kunal Verma, "METEOR-S Web Services annotation framework", In proceedings of the 13th international conference on WWW, ACM Press, 2004
- [25] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web", Scientific American Magazine, vol. 284, no. 5, pp. 34–43, 2001.
- [26] Derek Wright, "Cheap cycles from the desktop to the dedicated cluster: combining opportunistic and dedicated scheduling with Condor", Conference on Linux Clusters: The HPC Revolution, June, 2001,
- [27] Klaus Krauter, Rajkumar Buyya and Muthucumaru Maheswaran, "A taxonomy and survey of grid resource management systems for distributed computing ", Software: Practice and Experience (SPE), ISSN: 0038-0644, Volume 32, Issue 2, Pages: 135-164, Wiley Press, USA, February 2002.
- [28] Web Ontology Language (OWL), W3C Recommendation 2004, online: http://www.w3.org/2004/OWL/, 2004 .
- [29] Protégé :open source ontology editor and knowledge-base framework, <u>http://protege.stanford.edu</u>.
- [30] N. Nisan, S. London, O. Regev, and N. Camiel, Globally Distributed computation over the Internet: The POPCORN project, International Conference on Distributed Computing Systems (ICDCS'98), May 26 - 29, 1998, Amsterdam, The Netherlands, IEEE CS Press, USA, 1998.
- [31] Y. Aridor, M. Factor, and A. Teperman, cJVM: a Single System Image of a JVM on a Cluster, Proceedings of the 29th International Conference on Parallel Processing (ICPP 99), September 1999, Fukushima, Japan, IEEE Computer Society Press, USA
 - [32] Alchemi Documentation: http://www.cloudbus.org/ ~alchemi/files/1.0.beta/docs/AlchemiManualv.1.0.htm