ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 7, July 2014

Coalition of Clouds using Grid services and standards

Suresh Kumar RG [1], Vinoth Kumar R [2]

Assistant Professor, Rajiv Gandhi College of Engineering and Technology, Puducherry, India [1]

Assistant Professor, Rajiv Gandhi College of Engineering and Technology, Puducherry, India [2]

Abstract: Cloud computing is a model for facilitating expedient, on-demand network access to a shared pool of computing resources (e.g., networks, storage, servers, services, and applications) that can be rapidly provisioned and released with minimal management effort or service provider communication. This cloud model endorses availability and is composed of five essential characteristics (On-demand self-service, Resource pooling, Broad network access, Measured Service, Rapid elasticity). Also research communities such as AMO (Atomic, Molecular and Optical) physics research, Biology, and Neuroscience are investigating the relevance of Clouds, with their strengths and weaknesses in scientific environments. In this paper we will illustrate that in scientific environments there are some certain areas where cloud services are supporting the challenging e-Science requirements for virtual communities, resource discovery and dynamic service, resource federation and identity, and access to data catalogues. The Grid technology has actively contributed to address some of these issues, thus we propose to reuse existing Cloud services with Grid technology, experiences and production services, including standardization. In this paper we will provide strategy of how to apprehend the multi-cloud associated deployments based on a survey of existing Grid technologies in context expanding it with lessons learned gained in scientific environments. The contribution spotlight on the areas of data, compute, security and information. We will also show possible benefits that scientists can gain by adopting proposed solutions in cloud-based deployments.

Keywords: AMO; Grid Computing; multi-cloud; Clouds; e-Science

I. INTRODUCTION

Scientific computations are no longer a domain only for High Energy Physics (HEP), computation biology or climate modeling. Nowadays it becomes an every-day tool used even in social sciences and humanities that have been traditionally not very compute intensive. With the increasing popularity of the computational methods also the demand for compute, data, and networked resources increases. With broadening application of the computer Supported methods we also see changes in the resources demanded and paradigms used. Some of them cannot be easily satisfied by HPC/HTC resources offered by the re-search computer centres (e. g. dynamic expansion of compute resources into the unlimited region). It is also not practical nor economically justified for the new communities to create their own computer centres or to perform over-provisioning of resources for rare peak times in the research life-cycle. Therefore existing computing centres consider outsourcing resources e. g. by increasingly using Clouds. By this we are partly motivated by the success of the commercial public clouds, and partly to profit from the effect of scale and consolidation of resources.

Furthermore, the paper is also motivated by the work within the e-Infrastructure Reflection Group (e-IRG) that puts it this way: While academia has thus far approached distributed computing infrastructures using the Grid, useful for its federating ability and collaborative approach, the introduction of the new "Cloud Computing" paradigm demands a fresh assessment and attitude [1]. Cloud Computing offers compute, data and networked resources based on the pay-per-use, on-demand, and prudent manner. In the current situation scientists want to, however, access different resources (HPC/HTC, Storage, Clouds) offered by different providers (research computer centres or even commercial clouds). This is currently hard to achieve, due to a massive amount of existing computing and data resources, proprietary interfaces, and In-compatible authentication and authorization systems existing in public and private cloud offerings. The problem is even harder to solve for the cross-boundaries research endeavors where representatives from different communities are active.

To some extent the situation is similar to the situation that lead in the past to emerging of Grids [2], when it was required to have seamless access to distributed resources via standardized interfaces across multiple administrative and geographical boundaries. In this paper we propose to capitalize on these experiences and to work towards an Easy-accessible federation of heterogeneous resources. Through such a federation, scientific virtual communities could actively collaborate their results and provide access to distributed resources. As we mentioned earlier, that Grid communities have already invested ample amount of research to solve similar kind of challenges, and to establish unified view and access methods to federated resources. The clouds are, in our opinion, on the verge of establishing federated clouds i. e., to overcome the boundaries between different cloud providers and deployments. Thus we believe that the current cloud solutions could profit from the Grid experiences by incorporating the well-established and production-ready solutions from compute, data, security, and information system areas. By considering its strengths, we propose to reuse a set of existing middleware technologies and standards when implementing or deploying cloud solutions. In this work our analysis is based upon the technologies represented within the European Middleware Initiative Project (EMI) [3]. The European Middleware Initiative is a close collaboration of the three major middleware providers, ARC [4], gLite [5] and UNICORE [6], and other specialized software providers like cache [7] working towards a unified set of Grid middleware distributions. Furthermore, we will also put an emphasis on the set of available standards in the Grid domain which could bring benefits in terms of offering an easy access to larger, heterogeneous, virtual resources.

After surveying related work in Section II, our analysis of Grid services and standards applicable to support the vision of federated clouds is given in Section III. Where we list the existing technologies offered by EMI compute, and related standards that can be used to federate cloud computing VM access interfaces. In a similar manner the following sections IV, V, VI will analyze data, information, and security areas. We conclude our paper in Section VII.

II.RELATED WORK

Since the Cloud computing paradigm emerges in the context of e-Science infrastructure (aka Distributed Computing Infrastructures (DCI)), many publications and reports provide high-level statements like “including specialized portals hiding the certificate complexities of Grids as well as the possibility of learning from the Cloud virtualization approach.” [8]. Another report [1] “Seeking a way to combine best aspects of the Grid and Cloud Computing requires resolving technological and policy level conflicts between two paradigms that have started based on very different requirements, use cases and basic assumptions.” Finally, statements like “Europe should also build its innovative advantage in key areas through reinforced e-Infrastructures and through the targeted development of innovation clusters in key fields. It should develop an EU-wide strategy on cloud computing notably for government and science” were also published [9]. Despite the existence of such high-level publications we can conclude that no clear roadmap and vision for cloud developments in science really exists today.

An effort of clarification of and distinction from clouds was undertaken by the Grid community. The Grid paradigm started to evolve more than a decade ago with a publication by Foster et al [2]. It is thus appropriate to revisit Fosters latest views on how Grid and clouds are related. In [10], Foster compares both approaches side-by-side using different model perspectives (i. e., business, architecture, resource management, data, programming, application, and security). We agree with conclusion that clouds and Grids share a lot commonality in their vision, architecture, and technology. But it provides only several general overviews and a few concrete examples in context based on a few existing distributed systems developments available (mostly Globus [11]). In contrast, this paper focuses thus more on different approaches and techniques that contributed significantly to the success of Grids i. e., information model, broker, attribute authority, etc. and where existing developments are available provided by the EMI. It makes thus sense to survey related studies for this contribution with a focus on those initiatives that aim to reuse concrete existing Grid technologies and standards in cloud development activities.

Lezzi et al, provides in [12] a description of how the VENUS-C cloud platform can be used with e-Science applications. Within this architecture platform Grid standards such as OGSA-BES and JSDL are re-used for job management. In contrast to this contribution, our paper extends this view to other approaches beyond job management covering also information, security, and others. StratusLab [13] is another project that develops cloud technology primarily targeting the scenarios for offering Grid interfaces to virtual machines. The major breakthrough in this project is the planned support of Grid-based authentication and authorization mechanisms that is in-line with the finding of this paper. It allows users with X.509 credentials to seamlessly launch and monitor virtual machines. Apart from security area, the StratusLab approach tackles also other areas we mentioned, but being based on rather proprietary interfaces. In contrast, this paper aims to not only propose existing developments in context, but also suggests open standards in various areas (e. g., accounting, information, etc.). Finally, a survey of related approaches would be not complete if we not question the overall approach of “developing new clouds” with existing technology that still needs to be extended to meet the cloud users need. Instead, it might be wise to think about a complementary approach meaning “Integration of existing grid layers with the on-demand delivery model typical of commercial Clouds” as highlighted in [14], that paper tackles the question how e-Infrastructure can emerge to satisfy the needs of the emerging European Strategy Forum on Research Infrastructure (ESFRI) projects. For example, Di Meglio et al. [15] compared Grid and cloud integration scenarios with a generic overview of the number of combinations through which resource centers can effectively combine and offer both paradigms together.

The paper also articulates the implications on several standards and, analyzed how much overhead on technologies in terms of costs and implementation is required to provision both Grid and cloud interfaces at the same time. In contrast, this paper rather focuses on the different technical areas that are faced during integration of Grids and clouds instead of high-level combinations possibilities.

III-COMPUTE SERVICES

In distributed systems compute service is a capability offering interfaces for accessing CPU cores, including the management of the running applications life-cycle. In clouds, generally, we consider compute area is a package of virtual machine management and monitoring, resource allocation, and scheduling capabilities. The VM Management layer hides the complexity of hypervisor interaction, which includes dispatching VM to the appropriate physical hosts, and releasing resources after use.

The current landscape of cloud solutions is shaped by increasing number of the open source cloud solutions and public offerings of this service. With these advancements, there are upcoming challenges for scientific communities when it comes to interact with multiple of these kinds of cloud sites. This is more evident with virtual communities, each of which may consist of cloud endpoints supporting nonstandard interfaces. In an ideal case, a user should be able to submit the VM execution request to any of the sites available within a community, without the need to account for special properties of the particular cloud deployment. Once we contemplate on a scenario of VM run across different cloud platforms, there from we see two significant requirements: resource requirements description language, and remote interface managing and monitoring the life-cycle of the VM.

Under the domain of Grid computing, there have been similar challenges of offering compute resources to the virtual organizations. To deal with such requirements, EMI adopts broker-based technologies where scientists run jobs on hybrid Grid platforms. Within the EMI portfolio, WMS (Workload Management System) [16] suite allows the distribution and management of jobs to multiple, disparate computing resources. In a likewise manner, we could extend and apply this solution to enable scientific virtual communities forming a federated inter-cloud infrastructure. This will not only enhances the adoption in cloud communities, but also help resource centres willing to integrate their Grid and clouds offerings more seamlessly.

In order to federate heterogeneous clouds, standard compliance significantly reduces much overhead of end user clients interacting with diversified cloud interfaces hosted on an infrastructure [17]. From the past experiences, Grid community faced interoperability challenges in the establishment of the production e-Infrastructures [18], [19], [20], [21] wherein sites expose proprietary job management interfaces. In which the implied use of standards came to fore as a resolution for integrating multiple computing sites part of pan-continental initiatives. By having the experiences from infrastructure integration concerns, we see Grid "compute" standards such as OGSA-BES [22], and JSDL [23], became a viable choice for service implementers. As a matter of fact, these standards have been derived and adopted by major production Grid middleware providers [24], [25], and [26] to cater job management and monitoring scenarios. The standardization process of these interfaces then assured that broad agreements have been achieved in terms of the functionality it covers. Inspired by this concept, we can thus treat cloud VM as a Grid Job, which implies VM execution through the Grid standard execution service. In our proposal we envision the adoption of Grid standard interfaces to the cloud VM management layer, will positively influence the initiatives towards the vision of realizing inter-cloud infrastructure services and unified client tools. This approach will also invite the Grid users, who are early clouds adopters, to launch their customized VMs as execution environments using standards supported Grid clients.

IV. DATA MANAGEMENT

Many experts envision data-oriented science and computing as the most important paradigm of the e-Science in the coming years [27]. In fact the data volumes collected by the scientists grow in sizes and complexity. As a result of a distributed process of data collection and creation, the data are naturally distributed. The heterogeneity and distribution of these resources, however, make it hard to efficiently access and analyse the data collected by different communities, in different storages, in different points in time (using various data formats and data access methods).

Current cloud storages like Amazon Simple Storage Service (S3) [28] or OpenStack Swift [29], they offer access to distributed data and some replication functionalities but all that is confined within one cloud infrastructure. From the scientific perspective, this is a huge drawback leading to known problems such as "science data silos" that are not conveniently accessible by the broader scientific community. Hence, to be able to cope with the large amounts of distributed data we need to provide the users with a means to access them in an abstract way which hides the underlying complexity. The users usually do not care where the data are stored, they want to simply access or actually extract information from the data.

This known fact led to establishment few Grid technologies that are able to create federated data storages or data Grids. The EMI Portfolio includes few such a solution. Probably the most prominent example of a storage technology in Grids is dCache [7]. dCache offers a transparent access to disk-and tertiary (tapes,etc.) storages and allows for migration and replication of the data between physical locations without affecting the end-user. All the technical details are hidden from the user behind well-defined data access interfaces Storage Resource Manager (SRM) [30], http, WebDAV.

Similar functionalities (aggregated view and access to distributed disk storages) are also offered by: Disk Pool Manager (DPM) [31], StoRM [32] all three offer Storage Resource Management interface. The role of SRM in typical Grid ecosystem is to support basic data-related operations (file transfers and management, directory manipulation tools, replication and archiving data on mass storage systems). These functions are performed in a transparent way, for instance SRM implements protocol negotiation which allows to make files accessible from one physical location to other without any user intervention.

Another important aspect of distributed data management is the metadata management. Grid also provides tools for that, in this regard EMI offers a flexible metadata manager: AMGA [33]. AMGA enables annotation of data items with key-value metadata pairs and search functionality. The annotation is performed out-of-storage, that it is easily possible to use AMGA at federation level to create and describe collection of data resources

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 7, July 2014

independent from their physical location.

Above-mentioned Grid services and standards help to "merge" all different kind of data resources located within different infrastructures, allowing to virtually federating the data. The way a cloud data federation could be established with help of Grid services and standards would be to simply add an integration layer between currently distributed cloud storages and an inter-cloud instance of dCache or AMGA providing a unifying view of all the data and metadata resources. The interfaces of Storage Networking Industry Association (SNIA) such as Cloud Data Management Interface (CDMI) [34] might be useful in this context. Secondly it would be advisable to foster the usage of data access standards: SRM is one of the examples that need to undergo an evolution to be simpler if it should be fully used in cloud setups. However, it remains an excellent basis to start from. To reduce the hurdles of accessing different storages such an interface is needed.

V.INFORMATION SERVICES

Information services in clouds are the set of components to discover and publish capabilities and usage of the underlying resources. The information exposed by such services consumed for scenarios such as service availability, service discovery, service monitoring, and resource usage. Existing open source cloud solutions such as OpenStack [35], Open-Nebula [36], Eucalyptus [37] individually provides their own mechanisms of exposing resource information. Considering these solutions deployed on different sites as a part of an infrastructure, it will pose a plethora of requirements to form a consistent structure of representing such information. In order to unite resource information per se and the process of aggregation and register, we may require developing a set of components to perform functions mentioned earlier.

Again it is possible to learn from Grid experience in this context, since Grids are spanned over distributed regions and already deployed using different middleware's that are combined to form a collection of resources emitting information in a consistent manner. Whether it be a HPC or HTC based resource, the high level infrastructure administrators conveniently fetch an overview of different types of resource information and its key capabilities, depending on the site policies they can also further obtain a fine grained usage and accounting information. For catering such requirements, EMI produced a set of information components, which are mature and developed over years, and used in global-scale production infrastructures such as EGI [38].

Complete and correct information model is a basis for any type of service registry and information service. In the above mentioned information components, we see an urgent need to place an information model for information exchange purposes. Clearly the back-end model is also required for usage service as a foundation for integrating and uniting accounting information in an open, research based infrastructures. Similarly, Grid community published GLUE 2 [39] information model standard and Usage Record specification [40]. GLUE 2 information model provides an enriched representation of compute, and storage resources. These both kinds of resources are also present in cloud environments. In terms of GLUE 2 application for clouds, we need to perform minor extensions in terms of concepts and attributes to introduce entities peculiar to Virtual computing environments.

In modern distributed clouds, where federation is paramount, the services and resources are behaving in transient manner: being added to and removed from the federated infrastructure dynamically. Hence an infrastructure service registry which must reliably depict the status of cloud service's information is required. The lack of common service index drives the need of new service registry to perform the look-up and discovery in a unified manner. However, open source cloud solutions are already equipped with local service registries, which are used for internal management respective specific to the cloud middleware. As such, considering an infrastructure composed of different cloud distributions, local registries will not be useful for combining heterogeneous entities. EMI as a unified Grid middleware distribution supports the federation of Grid services through a common service registry called EMIR [41]. EMIR has been designed to index Grid services; furthermore it supports mechanism to federate distributed registry nodes in a robust manner. It supports a standardized GLUE 2 model to represent the service endpoints and offers a lightweight RESTful API to make registrations and query. In order to share usage information among multiple clouds, the sites need to communicate its resource accounting and usage information in standard format across the range of community based infrastructures. In the domain of Grids, OGF produced Usage Record (UR) specification [40] encapsulates the thorough semantics of accounting information model specialized for Grid resources. Specifically, in EMI we extended Usage Record specification with an emphasis on support of virtual

communities, and the usage of storage services. Along with OGF UR specification and EMI extensions we propose, are very much applicable to realize federated accounting infrastructure for heterogeneous and distributed cloud resources.

VI. SECURITY

Security is probably the hardest and most important problem that needs to be solved on the way towards cloud federation. It is a cross-cutting concern influencing all other aspects of the federation. Current situation in which each infrastructure provider use different authentication and authorization solutions constitutes a major hurdle for a seamless easy access to distributed resources. Long year effort of the Grid community lead to (almost) standardized way of accessing Grid resources. Grid uses certificate-based access to provide high-security on one hand and usage of resources offered by different providers on the other (credentials are not bound to a particular infrastructure).

Particularly in conduct of a harmonization effort EMI decided to use VOMS [42] as a product representing an Attribute Authority (AA) which releases signed security credentials with information beyond pure identities. VOMS authenticate user based on her Grid Certificate and issues X.509 attribute certificates and, more recently, Security Assertion Markup Language (SAML) authentication tokens. Second phase of accessing Grid resources is to derive authorization decisions. Within the EMI landscape this can be done by ARGUS [43]. It is responsible for generating consistent authorization decisions across geographically distributed Grid services. The internal ARGUS policies are described in a standardized way with help of eXtensible Access Control Markup Language (XACML) [44]. The interesting aspect of this security solutions proliferating in Grid is the notion of trust delegation. It allows authorizing access to data and computing resources not only for human users but also for other services acting on behalf of an end-user.

Nevertheless the final report of the 7th e-Infrastructure concertation meeting indicated clearly that certificates are complex too and the use of clouds can be the answer to "Find a way to alleviate certificates" [8]. It is a safe bet to say that the certificate-based authentication will not gain broad acceptance as a cloud access method. Thus the Grid services are not directly applicable as a building block for cloud federation. Some integration effort need to be undertaken in order to provide unified authentication and authorization methods. It is hard to assess the exact amount of work that need to be done to achieve such integration. Pragmatically in many cases it is not trivial at all. The commercial cloud solution providers might be reluctant to undertake any modification of their software. Although, even commercial providers can profit from the unification of security solutions, this may open opportunities for new communities. Most of the available open-source cloud solutions follow a modular design allowing extending or substituting infrastructure components. In fact OpenStack extracted their AA and created a stand-alone product [45], this can be interpreted as an "expression of will" to work on flexible authentication and authorization.

Also EMI become aware of the limitations of the certificate-based authentication in some cases. Thus an effort towards the use of more standardized interfaces in this filed has been undertaken, leading to a Security Token Service (STS) library. This can be used as a trusted source of security tokens derived from yet other tokens. STS basically translates "everything to everything", allowing to convert results of any given authentication method to any given authentication token. Clearly STS could be used to integrate clouds infrastructures with common authorization service (on federation level).

VII. CONCLUSION

In this paper I convey a vision of federated clouds and envisioned distributed compute and data resources unified, and providing easy, transparent access to the research users. I also pointed out that the vision could be realized by using Grid services and standards to bridge the gaps between current cloud offerings (be it commercial or private research clouds). Throughout the discussion I try to avoid to make a claim that one technology (Grid or cloud) is better than the other since there is much potential for being complementary to each other. Grid is around for many years thus it indeed provides some useful, production-ready solutions. So I provided the reader with some pointers showing applicable services and standards, and benefits of using them.

On the other hand, I also identified gaps in the federated clouds which cannot be bridged by just the integration of existing services and standards but rather require a common effort to extend them, and also to shape emerging use cases into the standardized interfaces and models.

REFERENCES

[1] e-IRG White Paper 2009, "http://www.e-irg.eu / images/stories/publ/white-papers/ erg_white_paper_2009_final.pdf.
[2] I. Foster, C.Kesselman, and S. Tuecke, "The anatomy of theGrid: Enabling scalable virtual organizations," International Journal of High Performance Computing Applications, vol. 15,pp. 200–222, 2001.
[3] "European Middelware Initiative," http://www.eu-emi.eu/.
[4] M. Ellert, et al., "Advanced resource connector middleware for lightweight computational grids," Future Generation Computer Systems, vol. 23, no. 2, pp. 219–240, 2007.
[5] E. Laure, et al., "Programming the Grid with gLite," in Computational Methods in Science and Technology, 2006, pp. 33–46 [6]
[6]A. Streit, et al., "Unicore 6 recent and future advancements,"Annals of Telecommunications, vol. 65, pp. 757– 762, 2010.
[7] P. Fuhrmann, "dCache: the commodity cache," in Proceedings of the 12th NASA Goddard and 21st IEEE Conference on Mass Storage Systems and Technologies, Apr. 2004, pp. 101–105.
[8] "Report of the 7th e-Infrastructure concertation meeting: e-Infrastructures for e-Science," http://tinyurl.com/7ovkcuw Oct.2009
[9] "A digital agenda for europe," European Commission, Tech. Rep.,2010
[10] . T. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared," in Proceedings of theGrid Computing Environments Workshop, Nov. 2008, pp. 1–10.
[11] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[12] I. Foster and C. Kesselman, "Globus: A metacomputing infrastructure toolkit," International Journal of Supercomputer Applications, vol. 11, pp. 115–128, 1996.
[13] "COMPSs in the VENUS-C Platform: enabling e-Science applications on the Cloud," in Proceedings of the 5th Iberian Grid Infrastructure Conference, Jun. 2011, pp. 73–84.
[14] "e-IRGBluePaper2010,"http://www.e-irg.eu/images/stories /eirg bluepaper 2010 final.pdf.
[15] A. Di Meglio, M. Riedel, S. Memon, C. Loomis, and D. Salomoni, "Grids and clouds integration and interoperability: an overview," in Proceedings of the International Symposium on Grids and Clouds and the Open Grid Forum, Mar. 2011, to appear
[16] M. Cecchi an det al. "The gLite workload management system," Journal of Physics: Conference Series, vol. 219, no. 6, pp. 1–12,2008.
[17] EDUCase Community clouds," http://tinyurl.com/6ruza4t
[18] C. E. Catlett, "TeraGrid: a foundation for US cyberinfrastructure," in Proceedings of the IFIP international conference on Network and Parallel Computing, Dec. 2005, pp.1–1.
[19] W. Gentzsch and et al., "DEISA: distributed europeaninfrastructure for supercomputing applications," Journal of Grid Computing, vol. 9, pp. 259–277, 2011.
[20] R. Pordes and et al., "The open science Grid," Journal of Physics:Conference Series, vol. 78, no. 1, pp. 1–15, 2007.
[21] F. Gagliardi and M.-E. Begin, "EGEE -providing a production quality grid for e-science," in Proceedings of the IEEE International Symposium on Mass Storage Systems and Technology, Jun. 2005, pp. 88–92.
[22] I. Foster, et al., "OGSA Basic Execution Service (BES), Version 1.0," OGF, Nov 2008.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 7, July 2014

[23] A. Anjomshoaa, et al., “Job Submission Description Language (JSDL) Specification, Version 1.0,” OGF, Jul. 2008.
[24] M. Memon and et al.., “Enhanced resource management capabilities using standardized job management and data access interfaces within UNICORE Grids,” in *Proceedings of the International Conference on Parallel and Distributed Systems*,Dec. 2007, pp. 1–6.
[25] E. Edelmann, et al. “Grid interoperation with ARC middleware for the CMS experiment,” *Journal of Physics: Conference Series*,vol. 219, no. 6, pp. 1–10, 2010.
[26] P. Andreetto, et al. “Standards-based job management in grid systems,” *Journal of Grid Computing*, vol. 8, pp. 19–45, 2010.
[27] J. Wood, et al., *Riding the wave: How Europe can gain from the rising tide of scientific data*. European Union, 2010, final report of the High Level Expert Group on Scientific Data: A submission to the European Commission.
[28] Amazon Simple Storage Service (S3), http://aws.amazon.com/s3
[29] “OpenStack Swift,” http://swift.openstack.org/.
[30] A. Sim, et al., “The Storage Resource Manager (SRM). Interface Specification, Version 2.2,” OGF, May 2008.
[31] G. Stewart, D. Cameron, G. Cowan, and G. McCance, “Storageand data management in EGEE,” in *Proceedings of the 5th Australasian symposium on ACSW frontiers*, Jan. 2007, pp. 69–77.
[32] A. Carbone, et al. “Performance studies of the StoRM storageresource manager,” in *Proceedings of the 3rd IEEE International Conference on e-Science and Grid Computing*, Dec. 2007, pp.423–430.
[33] B. Koblitz, N. Santos, and V. Pose, “The AMGA metadata service,” *Journal of Grid Computing*, vol. 6, pp. 61–76, 2008.
[34] “Cloud Data Management Interface Version 1.0,” SNIA, 2010.
[35] D. Milojii, I. M. Llorente, and R. S. Montero, “OpenNebula: A cloud management tool,” *IEEE Journal of Internet Computing*, vol. 15, pp. 11–14, 2011.
[36] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman,[37] L. Youseff, and D. Zagorodnov, “The Eucalyptus Open-Source Cloud-Computing System,” in *Proceedings of the IEEE/ACM Symposium on Cluster Computing and the Grid*, May 2009, pp.124–131.
[38] D. Kranzlm¨uller, J. Marco de Lucas, and P. Oster, *The European Grid* Initiative (EGI): Towards Sustainable Grid Infrastructure.Springer US, 2010, pp. 61–66.
[39] S. Andreozzi, et al., “GLUE Specification v. 2.0,” OGF, March 2009.
[40] R. Mach, et al., “Usage Record Format Recommendation,” OGF, 2007.
[41] “EMI Registry Requirements Specification and Architecture V.0.2,”http://cdsweb.cern.ch/record/1359909/files/EMIRegistryDesignv0.2.doc, EMI.
[42] R. Alfieri, R. Cecchini, V. Ciaschini, L. dell Agnello, A. Frohner,A. Gianoli, K. L¨orentey, and F. Spataro, “VOMS: an authorization system for virtual organizations,” *Lecture Notes in Computer Science*, vol. 2970, pp. 33–40, 2003
[43] “ARGUS Authorization Framework,” https://twiki.cern.ch/twiki/bin/view/EGEE/AuthorizationFramework.
[44] T. Moses and et al., “eXtensible Access Control Markup Language (XACML) Version 2.0,” OASIS, Feb. 2005.
[45] “OpenStack Keystone,” http://keystone.openstack.org/.