Editorial

Fabrizio Messina*

Department of Mathematics and Informatics,
University of Catania,
V.le Andrea Doria 6, 95125 Catania, Italy
Email: messina@dmi.unict.it
*Corresponding author

Rao Mikkilineni

C3DNA Inc.,
2520 Mission College Blvd, Suite 110,
Santa Clara, California 95054, USA
Email: rao@c3dna.com

Giovanni Morana

C3DNA Inc.,
Catania, Italy
Email: giovanni@c3dna.com

1 Introduction

Cloud and grid computing focus on the field of distributed service computing in various business and scientific areas. Virtualisation technology, commodity hardware, and high-bandwidth networks allow ubiquitous devices to access computational resources by means of a pay-as-you-go model. The demand for computing resources by public cloud service providers is ever increasing from users, e-service operators and scientific institutions, as the number and size of data centres distributed around the world is also increasing, such that complex services can be offered to all the customers around the world.

The use of cloud resources to reduce over-provisioning is the de facto standard to meet large fluctuations of workloads, while scientific organisations make use of the pay-per-use model to increase the capabilities of computing resources whenever is needed by business partners and scientists. As there is an ever increasing of high availability, performance and security of the services in highly distributed environments (Giunta et al., 2012; Messina et al., 2012; Giunta et al., 2013), recent cloud technologies focus on infrastructure management, inter-cloud and intra-cloud virtual image mobility, service composition, knowledge sharing. Moreover, the complexity of service creation, delivery and assurance processes is increasing. For instance online social networks (Messina et al., 2013b; Messina et al., 2013a) and service transaction management processes (Limthanmaphon and Zhang, 2004) are deployed across resources distributed at large scale by different cloud service providers.

Figure 1 shows how orchestrators and application management systems interact with infrastructure management systems. This growing complexity depends on the need to assure quality of service of applications at run-time in presence of an increasing number of servers and service providers. Many efforts are spent to improve reliability, availability, performance and security, i.e. to improve automation for such key concerns. As a consequence, there is an increasing of costs for development, deploying and management of applications.

Therefore providers have the need to *embed intelligence into cloud and grid infrastructures*, with particular focus on automation aimed at supporting applications through highly scalable patterns. Most of the efforts are focused on the integration of intelligence on resource managers, which operate at infra-structural level. A promising trend is represented by reliability to applications and services deployed on commodity infrastructure just as cellular organisms do. At this regards, John Von Neumann pointed out that *the basic principle of dealing with malfunctions in nature is to make their effects as unimportant as possible and to apply correctives, if they are necessary at all, at leisure. In our dealings with artificial automata, on the other hand, we require an immediate diagnosis. Therefore, we are trying to arrange the automata in such a manner that errors will become as conspicuous as possible, and intervention and correction follow immediately* (Aspray and Burks, 1986). There is an opportunity to discover post-hypervisor computing models and architectures, to decouple service management from infrastructure management systems at run-time and assure end-to-end distributed service transaction safety and survival. At the same time we think that the current complexity cliff must be avoided as much as possible: as Cockshott et al. (2012) observed, *the key property of general-purpose computer is that they are general purpose. We can use them to deterministically model any physical system, of which they are not themselves a part, to an arbitrary degree of accuracy. Their logical limits arise when we try to get them to model a part of the world that includes themselves.*

Any new computing models, management models (Messina et al., 2014a; Messina et al., 2014b) and programming models (Messina et al., 2007; Messina et al., 2013d; Messina et al., 2013c) aimed at integrating computations must support concurrency, mobility and synchronisation in order to execute distributed processes with global policy based management and local control. We must find ways to cross the Turing barrier (Turing, 1936; Mikkilineni, 2012) to deal with dynamic distributed processes (concurrency, synchronisation) at large scale. This will lead to bring together development and operations (DevOps) with a new approach to integrate functional and non-functional requirements of dynamic process management.

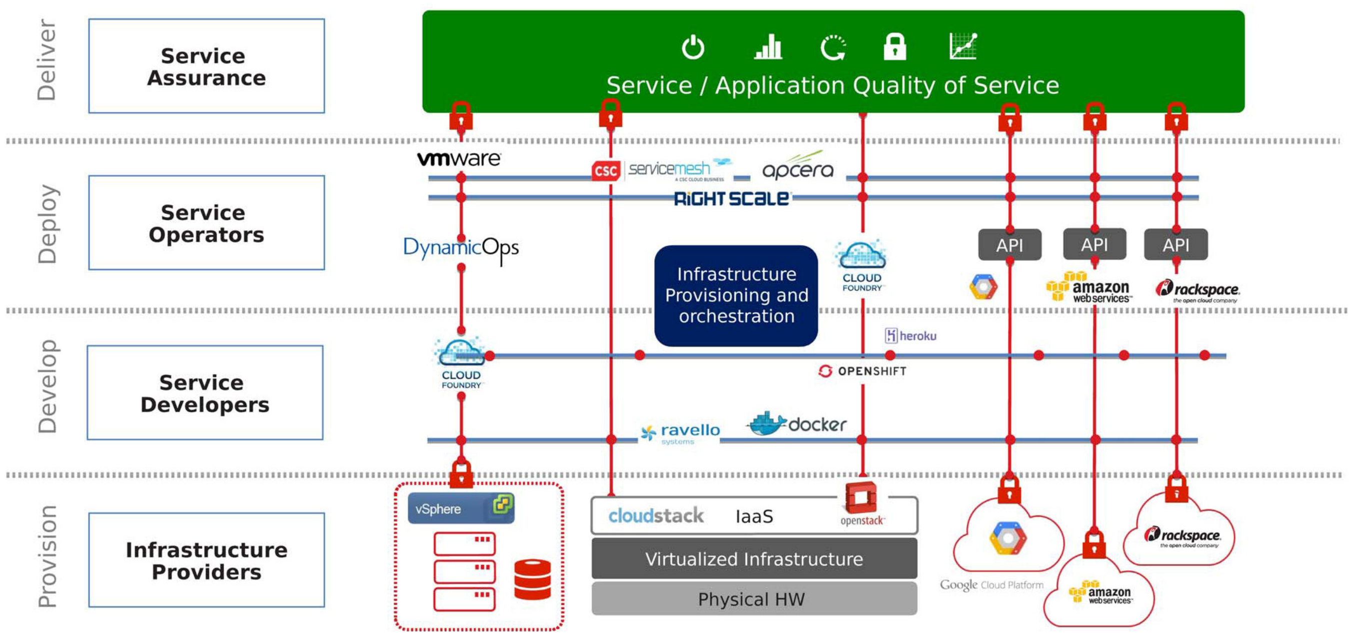
With WETICE-2009 we started to *analyse current trends in cloud computing, identify long-term research themes and facilitate collaboration in future research in the field that are not dictated or driven by the prototypical short term profit driven motives of a particular corporate entity* (Mikkilineni, 2012; Mikkilineni and Morana, 2013; Mikkilineni and Morana, 2012; Mikkilineni and Morana, 2011; Sarathy et al., 2010; Mikkilineni, 2009).

The 2014 edition of the track on the *Convergence of Distributed Clouds, Grids and their Management* embraced the need to develop and deploy *intelligent systems* to support new programming models and technologies for distributed grid and cloud computing centres at large scale. We focused on emerging infrastructures, middlewares, framework to support unified and federated utility computing infrastructures by embedding intelligent systems and innovative solutions to support the integration of grid and cloud systems. This led us to bring several interesting contributions on several fields, which have been expanded through further investigations. These works will be briefly discussed in the next section.

2 Summary of the contributions

Seven papers were selected owing to their quality and interesting topics concerning management of cloud and grid computing systems, i.e. complexity reduction by means of intelligent systems, algorithmic and scalable infrastructure solutions. Some of them address issues related to the management of the infrastructure, e.g. cognitive distributed computing or VM clustering. Intelligent, autonomous systems have been adopted for SLA negotiation and grid scheduling. Scientific computing has been taken into account, especially to bridge the gap between state-of-the-art of cloud infrastructures for scientific organisations and common, consolidated programming practices. The problems arising from the huge amount of information collected into cloud data centres, e.g. clustering, have been also addressed. In the following we provide a short summary of each contribution, to highlight the significance of the work.

Figure 1 The complexity of application quality of service assurance using inter-cloud and intra-cloud virtual machine image management



A novel near-parallel version of k-means algorithm for n-dimensional data objects using MPI deals with the problem of the exponential growing of data, which leads to huge bunches of information of different nature. It is presented an investigation on the use of fast techniques to discover and extract information from data pools. In particular, the main advantages of the k-means algorithm, which clusters data according to its main characteristics, are discussed. A fully parallel version for 1-dimensional objects, and in addition, a near-parallel approach for n-dimensional objects, is presented and discussed. The experimental results are interesting. Indeed, results studied for 1-dimensional objects are in-line with the theoretical outcomes, therefore, the authors have proved the correctness and the effectiveness of the developed technique. Moreover, experimental results for n-dimensional objects are very close with the outcome of the original sequential k-means, such that the authors could accept them as they are. The authors also propose to develop, in a future work, a variation of this algorithm in order to apply on non-numerical data, i.e. text, Boolean, raw text like names.

The authors of *Cognitive distributed computing: a new approach to distributed data centres with self-managing services on commodity hardware* continue to investigate how developing new computing models can bring a further contribution in reducing the complexity and create post-hypervisor inter-cloud and intra-cloud application and service transaction management. The main addressed problem is that there is a growing recognition that current infrastructure-centric deployment and management of web-scale applications is not able to scale and meet the demands of changing workloads and increasing demands on application availability, mobility and security. As a consequence, the authors have developed new management and programming models with a novel distributed intelligent managed element (DIME) network architecture to provide end to end service visibility and control across distributed utility computing infrastructures. The result is very interesting because their approach allows them to decouple application and service transaction management from myriad distributed infrastructure management and virtual machine movement at run-time, enabling policy based, secure-service mobility across physical servers or virtual machines deployed in data centres or public clouds. They show the advantage of using the new architecture by discussing an implementation of a web-stack provided with self-repair, auto-scaling, live-migration and end-to-end service transaction security. The authors finally remark that *the ability to dynamically change processes and the relationships between them is an important capability for enabling enterprises to function in the hyper-competitive, highly dynamic, on-demand business environment of today*, and that *the new approach using DNA unifies managing the life of a computation using the vital sign abstractions in the Cognitive Container and managing the intent of computations using the Cognitive Container networks and their management*.

In the paper entitled *A multi-agent protocol for service-level agreement negotiation in cloud federations*, the authors dealt with the need of assisting federated cloud providers when negotiating SLA across different sites at large scale. In particular, they designed a multi-agent SLA negotiation protocol which does not need a common global agent ontology, because the protocol takes into account different concerns related to the understanding of semantic and technical terms, and the knowledge about the agent abilities on semantic and technical terms are shared among the cloud federation. In order to show the advantages of the proposed approach, they discuss an example of the protocol.

As grid computing is nowadays used in many scientific organisations (also in conjunction with public cloud computing centres), the authors of *A hybrid particle swarm optimisation-genetic algorithm applied to grid scheduling* have been exploring the scheduling problems in order to make efficient use of the available grid resources, which allows to limit the use of external (cloud) resources, thus allowing scientific organisations to have a higher level of autonomy. In particular, the work focuses on the application of the particle swarm optimisation (PSO) meta-heuristic to the scheduling of independent users' jobs on grids, showing that the PSO method can achieve satisfactory results in simple problem instances, but it has a tendency to stagnate around local minima in high-dimensional problems. Therefore the authors propose a novel hybrid particle swarm optimisation-genetic algorithm (H_PSO) method in order to increase swarm diversity when a stagnation condition is detected. They have shown that H_PSO can successfully improve the scheduling solution.

Another interesting work in the field of scientific computing is entitled *A framework for cloud-aware development of bag-of-tasks scientific applications*, on which the authors give a contribution to exploit the potential of cloud computing for scientific applications. They designed and developed a friendly environment, able to exploit cloud scalability and to provide fault tolerance. In particular, they present an interesting framework built on the top of a cloud-aware platform (mOSAIC) for the development of bag-of-tasks scientific applications. Their contribution is different from other work as the resulting solution makes it fairly easy for a scientist to develop and to run applications on the cloud with scalability and fault-tolerance provided out-of-the box by the programming framework. Finally, they point out that communication, synchronisation, scalability and fault- tolerance are provided by the framework and by the underlying platform, so they do not present an issue to the developer.

The authors of the paper *DynamoGraph: extending the Pregel paradigm for large-scale temporal graph processing* focus on the interesting applications of large-scale temporal graphs, which can serve as a model in many scientific scenarios, like studies on online social networks, research interest in gathering and analysing data about human behaviour and interaction. Temporal graphs gain attraction in the analysis of dynamic processes in social networks. In their work a cloud-based distributed processing framework extending the Pregel paradigm for large-scale temporal graphs is presented. The work is motivated by the fact that using computing resources in the cloud allows applications designed for analysis to be scalable and prepared to scale for the massive datasets that occur, e.g. in social network analysis. Their contribution allows scientists to use a Pregel-style processing paradigm to process large-scale social network data without incurring into limitations like, e.g. lack support for temporal graphs, or support datasets of limited size only.

Authors of *A comparison of techniques to detect similarities in cloud virtual machines* deal with scalability in monitoring and management of cloud data centres by a novel approach that focuses on the clustering of virtual machines (VMs) exhibiting similar behaviour. They studied existing solutions for automatic VM clustering, stating that they present some important drawbacks that hinder their applicability to real cloud scenarios. They start from two considerations. The first one is that existing solutions show a trade-off between the accuracy of the VMs clustering and the computational cost of the automatic process. The latter is represented by the fact that their performance shows a strong dependence on specific technique parameters. Their approach is designed to overcome the issues above. They make use of a Mixture of Gaussians (MoGs) together with the Kullback Leiber divergence to model similarity between VMs. They also provide a thorough experimental evaluation of the proposal and of existing techniques to identify the most suitable solution for different workload scenarios.

3 Conclusions

From the 2014 edition, the track on the Convergence of Distributed Clouds, Grids and their Management exploits a trend aimed at developing and deploying intelligent systems for distributed grid and cloud at large scale. Authors mostly focused on infrastructure management to support the composition of unified and federated utility computing infrastructures by embedding intelligent systems and smart solutions. This issue includes several interesting contributions focusing on the management of cloud and grid computing systems, especially for complexity reduction. Some interesting works concern cognitive distributed computing and intelligent, autonomous systems. We finally point out that scientific computing has received much interest, with the aim of bridging the gap between state-of-the-art of the support of cloud infrastructures for scientific organisations and consolidated programming practices.

References

Aspray, W. and Burks, A. (1986) *Papers of John Von Neumann on Computing and Computer Theory*, MIT Press Cambridge, MA, USA.

Cockshott, P., Mackenzie, L.M. and Michaelson, G. (2012) *Computation and its Limits*, Oxford University Press.

Giunta, R., Messina, F., Pappalardo, G. and Tramontana, E. (2012) 'Augmenting a web server with QoS by means of an aspect-oriented architecture', *IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.179–184.

Giunta, R., Messina, F., Pappalardo, G. and Tramontana, E. (2013) 'Kaqudai: a dependable web infrastructure made out of existing components', *IEEE 22nd International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.146–151.

Limthanmaphon, B. and Zhang, Y. (2004) 'Web service composition transaction management', *Proceedings of the 15th Australasian Database Conference*, Vol. 27, Australian Computer Society, Inc., pp.171–179.

Messina, F., Pappalardo, G. and Santoro, C. (2012) 'Decentralised resource finding in cloud/grid computing environments: a performance evaluation', *IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.143–148.

Messina, F., Pappalardo, G. and Santoro, C. (2013c) 'Complexsim: a flexible simulation platform for complex systems', *International Journal of Simulation and Process Modelling*, Vol. 8, No. 4, pp.202–211.

Messina, F., Pappalardo, G. and Santoro, C. (2013d) 'Integrating cloud services in behaviour programming for autonomous robots', *Algorithms and Architectures for Parallel Processing*, Springer, pp.295–302.

Messina, F., Pappalardo, G. and Tramontana, E. (2007) 'Design and evaluation of a high-level grid communication infrastructure', *Concurrency and Computation: Practice and Experience*, Vol. 19, No. 9, pp.1299–1316.

Messina, F., Pappalardo, G., Rosaci, D. and Sarné, G.M. (2014a) 'A trust-based, multi-agent architecture supporting inter-cloud VM migration in IaaS federations', *Internet and Distributed Computing Systems*, Springer International Publishing, pp.74–83.

Messina, F., Pappalardo, G., Rosaci, D. and Sarné, M.G. (2014b) 'An agent based architecture for VM software tracking in cloud federations', *8th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)*, IEEE, pp.463–468.

Messina, F., Pappalardo, G., Rosaci, D., Santoro, C. and Sarné, G.M.L. (2013a) 'A distributed agent-based approach for supporting group formation in P2P e-learning', *AI* IA 2013: Advances in Artificial Intelligence*, Volume 8249 of *Lecture Notes in Computer Science*, Springer, pp.312–323.

Messina, F., Pappalardo, G., Rosaci, D., Santoro, C. and Sarné, G.M.L. (2013b) 'HySoN: a distributed agent-based protocol for group formation in online social networks', *Multiagent System Technologies*, Springer, pp.320–333.

Mikkilineni, R. (2009) 'WETICE 2009 first workshop on collaboration and cloud computing', *18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2009. WETICE'09*, pp.28–32.

Mikkilineni, R. (2012) 'Going beyond computation and its limits: injecting cognition into computing', *Applied Mathematics*, Vol. 3, No. 11A, pp.1826–1835.

Mikkilineni, R. and Morana, G. (2011) 'Convergence of distributed clouds, grids and their management track report – CDCGM 2011', *20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.94–97.

Mikkilineni, R. and Morana, G. (2012) 'CDCGM track report: convergence of distributed clouds, grids and their management', *IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.121–124.

Mikkilineni, R. and Morana, G. (2013) 'CDCGM track report', *IEEE 22nd International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp.131–133.

Sarathy, V., Narayan, P. and Mikkilineni, R. (2010) 'Next generation cloud computing architecture: enabling real-time dynamism for shared distributed physical infrastructure', *19th IEEE International Workshop on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)*, IEEE, pp.48–53.

Turing, A.M. (1936) 'On computable numbers, with an application to the entscheidungsproblem', *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 42.