CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
*Concurrency Computat.: Pract. Exper.* (2011)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.1865

EDITORIAL

# Special section on autonomic cloud computing: technologies, services, and applications

Welcome to the special issue of *Concurrency and Computation: Practice and Experience* (CCPE) journal. This special issue compiles a number of excellent technical contributions that significantly advance the state-of-the-art in autonomic cloud computing.

Cloud computing [1, 2] is an emerging utility computing model that allows users to dynamically access, select, and configure a large pool of IT resources (virtual machine templates, storage, and networking elements) and deliver them as 'computing utilities' to consumers in a pay-as-you-go manner. Several vendors have emerged in this space including IBM, VMware, Microsoft, Manjrasoft, and Yahoo. This model of computing is quite attractive, especially for small and medium sized enterprises, as it allows them to focus on consuming or offering services on top of cloud infrastructure. At high-level, cloud computing might not seem radically different from the existing paradigms: World Wide Web, grid computing, and cluster computing. However, key differentiators of cloud computing are its technical characteristics such as on-demand resource pooling or rapid elasticity, self-service, almost infinite scalability, end-to-end virtualization support, and robust support of resource usage metering and billing. Additionally, nontechnical differentiators include services that are offered under pay-as-you-go-model, guaranteed Service Level Agreement (SLA), faster time to deployments, lower upfront costs, little or no maintenance overhead, and environment friendliness.

Unpredictability is a fact in a distributed computing environment, and the Cloud is no exception. Performance unpredictability [3] in the Cloud is in fact a major issue for many users and it is coined as one of the major obstacles for cloud computing. For instance, researchers (biologists, physicists, finance analysts, etc.) expect guaranteed performance for their experiments, independent of the current workload and state [4] of IT resources of the Cloud, because this is key to repeatability of results. Other examples are small and medium sized enterprises (gaming company, web application providers) that want strict assurance on SLA; for example, an end-user request for a web page or multimedia content has to be served within the agreed time-limit. Hence, it is highly important for Cloud vendors that they have the ability to offer guaranteed SLAs based on performance metrics — such as response time and throughput. Interestingly, vendors seem to base their SLAs on availability of their offering, while completely ignoring response time and throughput. Hence, it is clear that dealing with performance unpredictability is critical to exploiting the full potential of clouds. In this special issue, we have tried to compile some high quality papers that exhaustively deal with some of the aforementioned issues. Next, we briefly describe the technical contributions, which were selected for publication in this special issue. All of the selected papers underwent a rigorous peer-review process.

The end-to-end QoS negotiation for SLA establishment for composite services involves compound multiparty negotiations in which the composite service provider concurrently negotiates with multiple candidates for each atomic service, selecting the one that best satisfies the atomic service QoS preferences while ensuring that the end-to-end QoS requirements are also fulfilled. It is necessary to derive the atomic utility boundaries from the global utility boundary to be able to negotiate with potential candidates. Additionally, there has to be a mechanism for updating these boundaries in subsequent negotiation rounds based upon the individual negotiation outcomes. To counter these complexities, in paper [5] titled 'Establishing Composite SLAs through Concurrent QoS Negotiation with Surplus Redistribution', Richter *et al.* propose an algorithm for the decomposition of global utility boundary into atomic service utility boundaries, and the surplus redistribution from successful negotiation outcomes among the remaining negotiations. The proposed mechanism is a practical approach to efficiently coordinate concurrent service negotiations within complex workflows, enabling the iterative and interactive adjustment of the negotiation boundaries for each atomic service in a composition based on the performance of other atomic negotiations. They demonstrate the feasibility of our approach by evaluating it with some popular negotiation strategies using the Specialised Property Search Scenario.

Many scientific workflows are data intensive where large volumes of intermediate data are generated during their execution. Some valuable intermediate data need to be stored for sharing or reuse. Traditionally, they are selectively stored according to the system storage capacity determined manually. As doing science in the Cloud has become popular nowadays, more intermediate data can be stored in scientific cloud workflows based on a pay-for-use model. In the paper in [6] titled 'A data dependency based strategy for intermediate data storage in scientific cloud workflow systems', Yuan *et al*. build an intermediate data dependency graph (IDG) from the data provenance in scientific workflows. With the IDG, deleted intermediate data can be regenerated, and as such they develop a novel intermediate data storage strategy that can reduce the cost of scientific cloud workflow systems by automatically storing appropriate intermediate data sets with one Cloud service provider. The strategy has significant research merits, that is, it achieves a cost-effective trade-off of computation cost and storage cost and is not strongly impacted by the forecasting inaccuracy of data sets' usages. Meanwhile, the strategy also takes the users' tolerance of data accessing delay into consideration. Authors utilize Amazon's cost model and apply the strategy to general random and specific astrophysics pulsar searching scientific workflows for evaluation. The results show that our strategy can reduce the overall cost of scientific cloud workflow execution significantly.

Recall that, one of the biggest premises of cloud computing is the flexibility of delivering IT resources and virtual appliances as an utility such as phone, electricity, gas, and water services. It enables users to have access to computing infrastructure, platform, and software as services over the Internet. To be competitive, however, Cloud providers need to be able to adapt to the dynamic loads from users, not only optimizing the local usage and costs but also engaging into agreements with other clouds to complement local capacity. The infrastructure in which competing clouds are able to cooperate to maximize their benefits is called a Federated Cloud. Just as clouds enable users to cope with unexpected demand loads, a Federated Cloud will enable individual clouds to cope with unforeseen variations of demand. The definition of the mechanism to ensure mutual benefits for the individual clouds composing the federation, however, is one of its main challenges. Gomes *et al*. in their paper [7] 'Pure exchange markets for resource sharing in federated clouds' propose and investigate the application of market-oriented mechanisms based on the General Equilibrium Theory of Microeconomics to coordinate the sharing of resources between the clouds in a Federated Cloud.

Several research institutions and universities own computational capacity that is not effectively utilized, thereby providing an opportunity for such institutions to use such capacity to offer Cloud services (to both internal and external users). However, the unreliability and unpredictability of these resources mean that their use in the context of an SLA is high risk, leading to a reduction in reputation and economic penalties in case of SLA violation. To overcome these challenges, in the paper [8] titled 'Towards autonomic management for Cloud services based upon volunteered resources', Caton and Rana propose a methodology that addresses the issues of unreliability and unpredictability such that Cloud software services could be hosted upon volunteered resources. To enable the harnessing of these resources, they rely on autonomic fault management techniques that allow such systems to independently adapt to the resources they use based upon their perception of individual resource reliability. Using the proposed approach they were able to scale out the backend infrastructure of the Cloud service elastically (minimum 30 s per worker), opportunistically, and autonomically. To summarize, the authors address two key questions in their paper: Can a campus volunteer infrastructure be used in Cloud provisioning? and What measures are necessary to ensure reliability at the resource level?

To improve the hosting and delivery of applications through cloud-based IT resources, Champrasert *et al*. in the paper [9] titled 'Exploring self-optimization and self-stabilization properties in bio-inspired autonomic cloud computing', describe architecture to build self-optimizable and self-stabilizable applications. The design of the proposed architecture, SymbioticSphere, is inspired by key biological principles such as decentralization, evolution, and symbiosis. In SymbioticSphere,



DOI: 10.1002/cpe

each cloud application consists of application services and middleware platforms. Each service and platform is designed as a biological entity, and implements biological behaviors such as energy exchange, migration, reproduction, and death. Each service/platform possesses behavior policies, as genes, each of which defines when and how to invoke a particular behavior. SymbioticSphere allows services and platforms to autonomously adapt to dynamic network conditions by optimizing their behavior policies with a multi-objective genetic algorithm. Moreover, SymbioticSphere allows services and platforms to autonomously seek stable adaptation decisions as equilibria (or symbiosis) between them with a game theoretic algorithm. This symbiosis augments evolutionary optimization to expedite the adaptation of agents and platforms. It also contributes to stable performance that contains a very limited amount of fluctuations. Simulation results demonstrate that agents and platforms successfully attain self-optimization and self-stabilization properties in their adaptation processes.

We hope that the readers will find the articles of this special issue to be informative and useful.

RAJIV RANJAN
*CSIRO Information and Communication Technologies (ICT) Centre*
*Information Engineering Laboratory, Computer and Information Technology Building*
*Australian National University, North Road, Acton, ACT 2601, Australia*

RAJKUMAR BUYYA
*Cloud Computing and Distributed Systems (CLOUDS) Laboratory*
*Department of Computer Science and Software Engineering*
*The University of Melbourne, Australia*

MANISH PARASHAR
*NSF Center for Autonomic Computing*
*Department of Electrical and Computer Engineering*
*Rutgers: The State University of New Jersey, U.S.A.*

## REFERENCES

1. Buyya R, Broberg J, Goscinski A (eds). *Cloud Computing: Principles and Paradigms*. Wiley Press: USA, 2011. 644 pages, ISBN-13:978-0470887998.
2. Wang L, Ranjan R, Chen J, Benatallah B. Cloud Computing: Methodology, Systems, and Applications. ISBN: 9781439856413, 844 Pages, CRC Press, Taylor and Francis Group, October 03, 2011.
3. Armbrust M *et al.* A view of Cloud Computing. In *Communications of the ACM Magazine*, Vol. 53, Issue 4. ACM Press: New York, USA, 2010; 50–58.
4. Schad J, Dittrich J, Quian J. Runtime measurements in the Cloud: Observing, Analyzing, and Reducing Variance. *Proceedings of VLDB Endowment* 2010; **3**(1-2):460–471.
5. Richter J, Chhetri MB, Kowalczyk R, Bao Vo Q. Establishing Composite SLAs through Concurrent QoS Negotiation with Surplus Redistribution. *Concurrency and Computation: Practice and Experience* DOI: 10.1002/cpe.1727.
6. Yuan D, Yang Y, Liu X, Zhang G, Chen J. A Data Dependency based Strategy for Intermediate Data Storage in Scientific Cloud Workflow Systems. *Concurrency and Computation: Practice and Experience* DOI: 10.1002/cpe.1636.
7. Gomes ER, Vo QB, Kowalczyk R. Pure Exchange Markets for Resource Sharing in Federated Clouds. *Concurrency and Computation: Practice and Experience* DOI: 10.1002/cpe.1659/.
8. Caton S, Rana O. Towards Autonomic Management for Cloud Services based upon Volunteered Resources. *Concurrency and Computation: Practice and Experience* DOI: 10.1002/cpe.1715.
9. Champrasert P, Suzuki J, Lee C. Exploring Self-Optimization and Self-Stabilization Properties in Bio-inspired Autonomic Cloud Applications. *Concurrency and Computation: Practice and Experience* DOI: 10.1002/cpe.1906.