UNIVERSITI PUTRA MALAYSIA

ADAPTABLE SERVICE LEVEL AGREEMENT EVALUATION FRAMEWORK BASED ON DYNAMIC MONITORING INTERVAL IN CLOUD COMPUTING

NOR SHAHIDA MOHD JAMAIL

FSKTM 2016 17

UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI

ADAPTABLE SERVICE LEVEL AGREEMENT EVALUATION FRAMEWORK BASED ON DYNAMIC MONITORING INTERVAL IN CLOUD COMPUTING

By

NOR SHAHIDA BINTI MOHD JAMAIL

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

August 2016

COPYRIGHT

DEDICATION

This thesis is especially dedicated to:

My Mother and Father,
Your love and enlightenment follow me no matter where I go.

My husband,
For all that you have been, for all that you are and will always be, I am grateful.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

ADAPTABLE SERVICE LEVEL AGREEMENT EVALUATION FRAMEWORK BASED ON DYNAMIC MONITORING INTERVAL IN CLOUD COMPUTING

By

NOR SHAHIDA BINTI MOHD JAMAIL

August 2016

Chair : Rodziah Atan, PhD

Faculty : Computer Science and Information Technology

Service level agreement (SLA) is a contract between service provider and consumer in relation to the quality of service (QoS) in cloud computing. SLA contains agreed attributes and value of service level objectives (SLO). Service provider should deliver services based on agreed quality in SLA. For example, QoS value of running service exceeds the agreed SLO, service provider ought to bear certain amount of penalty or compound fees for the SLA violation. SLA monitoring tool is unavoidable to assess the agreed SLAs at run time and detect any probable SLA violations. Both service provider and consumer need to monitor and assess QoS to ensure SLA validity.

The cost value and benefit value of SLA monitoring systems is a concerned issue in cloud computing. The SLA monitoring systems need resources such as CPU, Memory, or execution storage. The amount of consumed resources by monitoring system is the cost of SLA evaluation. On the other hand, SLA monitoring systems make benefits by detecting SLA violations in a sense that service provider subsequently can adapt the infrastructure to prevent more numbers of SLA violations and avoid penalty cost. The interval value of monitoring system has a direct impact to the cost and benefit values of monitoring system. Experiment results for CASViD and LoM2HiS frameworks have demonstrated that short measurement intervals negatively affect the overall system performance, whereas long measurement intervals cause heavy undetected SLA violations. Current monitoring systems have often constant interval value and individual monitoring process for all services and SLAs which it decreases the adaptability and cannot make a balance between cost value and benefit value of SLA monitoring systems. Subsequently, service providers face either high monitoring overhead or significant number of undetected SLA violations.

This study proposes SLA monitoring framework to evaluate the agreed SLA and detect SLA violations in cloud computing. The proposed framework locates a trusted-party between service provider and service consumer to maintain and evaluate SLAs, and without any manipulation of permission from service provider, the monitoring engine validate the SLA independently for each SLA data collector thread. This study also proposed a dynamic monitoring interval (DMI) to make a balance between cost values and benefit values of SLA monitoring system in different environmental situations (ES) monitored at run time.

The experimental testbed is established with one server and four virtual machines (VMs). The developed monitoring tools evaluate the predefined SLA of a selected experiment case and produced results demonstrating that the proposed DMI have the highest profit in normal ES by 23.6% growth compared to the best static interval applied to the other four VMs. Each constant interval gained a high profit in specific situation and had a low profit in other ESs whereas the proposed DMI adapted the interval value successfully and achieved the high profit in all ESs compared to the static intervals.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENYESUAIAN RANGKA KERJA PENILAIAN DALAM PENGAJIAN TAHAP PERKHIDMATAN BERDASARKAN PADA SELANG PEMANTAUAN DINAMIK DALAM PERKOMPUTERAN AWAN

Oleh

NOR SHAHIDA BINTI MOHD JAMAIL

OGOS 2016

Pengerusi : Rodziah Atan, PhD

Fakulti : Sains Komputer dan Teknologi Maklumat

Perjanjian tahap perkhidmatan (SLA) adalah kontrak antara pembekal perkhidmatan dan pengguna berhubung dengan kualiti perkhidmatan (QoS) dalam perkomputeran awan. SLA mengandungi sifat-sifat bersetuju dan nilai objektif tahap perkhidmatan (SLO). Pembekal perkhidmatan perlu menyampaikan perkhidmatan berdasarkan kualiti yang dipersetujui dalam SLA. Sebagai contoh, nilai QoS perkhidmatan yang diberikan melebihi SLO yang telah dipersetujui, pembekal perkhidmatan patut menanggung sejumlah bayaran penalti atau kompaun kerana melanggar SLA. Alat pemantauan SLA tidak dapat menilai SLA yang dipersetujui pada masa pengoperasian dan mengesan sebarang pelanggaran SLA yang berkemungkinan. Kedua-dua pembekal perkhidmatan dan pengguna perlu memantau dan menilai QoS untuk memastikan kesahihan SLA.

Nilai kos dan nilai faedah sistem pemantauan SLA adalah isu yang diambil berat dalam perkomputeran awan. Sistem pemantauan SLA memerlukan sumber seperti CPU, memori, atau penyimpanan data. Jumlah sumber yang digunakan oleh sistem pemantauan adalah kos penilaian untuk SLA. Sebaliknya, sistem pemantauan SLA mendapat faedah dengan mengesan pelanggaran SLA dalam erti kata bahawa pembekal perkhidmatan akan dapat menyesuaikan diri dengan infrastruktur untuk mencegah lebih banyak nombor pelanggaran SLA dan mengelakkan kos penalti. Nilai selang bagi sistem pemantauan mempunyai kesan langsung kepada kos dan faedah terhadap sistem pemantauan tersebut. Keputusan pemerhatian untuk rangka kerja CASViD dan LoM2HiS telah menunjukkan bahawa selang ukuran pendek sangat negative dan menjejaskan prestasi keseluruhan sistem, manakala selang ukuran panjang menyebabkan pelanggaran SLA yang tidak dapat dikesan dengan banyak. Sistem pemantauan sedia ada mempunyai nilai selang yang tetap dan proses pemantauan individu

untuk semua perkhidmatan dan SLA dimana mengurangkan keupayaan menyesuaikan diri dan tidak boleh membuat keseimbangan antara nilai kos dan nilai faedah untuk sistem pemantauan SLA. Selain itu, pembekal perkhidmatan berdepan dengan pemantauan yang terlalu tinggi atau sejumlah besar pelanggaran SLA tidak dapat dikesan.

Kajian ini mencadangkan rangka kerja pemantauan SLA untuk menilai SLA yang dipersetujui dan mengesan pelanggaran SLA dalam perkomputeran awan. Rangka kerja yang dicadangkan menempatkan pihak dipercayai antara pembekal perkhidmatan dan pengguna perkhidmatan untuk menyelenggara dan menilai SLA, dan tanpa apa-apa manipulasi daripada pembekal perkhidmatan, enjin pemantauan mengesahkan SLA bebas bagi setiap pengumpulan data SLA. Kajian ini juga mencadangkan selang pemantauan dinamik (DMI) untuk membuat keseimbangan antara nilai kos dan nilai faedah untuk sistem pemantauan SLA berada dalam keadaan persekitaran yang berbeza (ES) yang dipantau pada masa perlaksanaan.

Ruang ujian ditubuhkan dengan satu pelayan dan empat mesin maya (VM). Alat pemantauan dibangunkan untuk menilai SLA yang dipratentukan kes percubaan terpilih dan menghasilkan keputusan yang menunjukkan bahawa DMI yang dicadangkan mempunyai keuntungan yang tertinggi di ES normal oleh pertumbuhan 23.6% berbanding dengan selang tetap yang diuji daripada empat VM. Setiap selang tetap mendapat keuntungan yang tinggi dalam keadaan tertentu dan mempunyai keuntungan yang rendah dalam ES lain sedangkan DMI yang dicadangkan disesuaikan dengan berjaya untuk nilai selang dan mencapai keuntungan yang tinggi dalam semua ES berbanding dengan jangka masa yang tetap.

ACKNOWLEDGEMENTS

In the name of Allah, the Beneficent, the Compassionate , and who always give me strength and patience to complete my duties through many challenges situations.

First of all, I would like to say Alhamdulillah, a lot of thanks to Allah S.W.T for His blessing and kindness and also peace and blessing to the Prophet Muhammad (peace be upon him) the messenger of Allah.

I was successful to complete my PhD project and this thesis as the symbol of the idea, support and guidance that I get from all my lecturers, family members and colleagues. Kindly regards also to my educational sponsor bodies, the Ministry of education Malaysia and Universiti Pendidikan Sultan Idris, Malaysia.

First and most importantly, I would like to express the deepest gratitude to my supervisor, Assoc. Prof. Dr. Rodziah Atan, for her supports regarding research and scholarship. Her continued advice, guidance and encouragement led me to the right way. I would like to extend my appreciation to other committee members, Professor Dr. Rusli Abdullah and Dr. Mar Yah Said, for their effective advices during my research. I am also very thankful to Faculty of Computer Science and Information Technology, UPM.

My deep appreciation also goes to my family members who were so tolerant and support me all these years. Special thanks to my husband, Wan Mohd Shahid b. Mohd Idros for his encouragement supports and trust that he gives to me. Besides, thanks to all my kids for their understanding and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AM	Autonomic Manager
CEP	Complex Event Processor
CPU	Central Processing Unit
CSLA	Cloud Service Level Agreement
DMI	Dynamic Monitoring Interval
DSB	Distributed Service Bus
ES	Environmental Situation
FI	Future Internet
ICT	Information and Communication Technology
OGF	Open Grid Forum
PC	Personal Computer
QOS	Quality Of Service
RBSLA	Rule-based Service Level Agreement
RU	Recommending users
SLA	Service Level Agreement
SLAM	SLA Management
SLO	Service Level Objective
SOA	Service Oriented Architecture
SP	Service Provider
TP	Third-party
UPM	Universiti Putra Malaysia
UPSI	Universiti Pendidikan Sultan Idris
VM	Virtual Machine
W3C	World Wide Web Consortium
WS	Web Service
WSLA	Web Service Level Agreement
WSML	Web Service Management layer
WSMO	Web Service Modeling Ontology
WSOL	Web Service Offering Language
XHTML	Extensible Hypertext Markup Language
XML	Extensible Markup Language

CHAPTER 1

INTRODUCTION

1.1 Background

Service Level Agreements (SLA) is used to define functional and non-functional cloud services to be agreed by both service customers and providers as presented in Figure 1.1 SLA in cloud computing (Rimal et al., 2011). Common parameters included in SLA are such as: pricing model, usage model, resource metering, billing, and monitoring (Buyya et al., 2010). When a service provider is unable to comply with terms stated in certain SLA, a violation occurs (Wu and Buyya, 2010). For example, an IaaS from cloud service provider may guarantee a specific response time of a VM, exact value of storage space or reliability of data. However, when a customer does not get the promised response time, virtual disk space or access failures, the SLA is violated. Therefore, SLA monitoring tools is unavoidable to evaluate agreed SLAs at the run time and detect any probable SLA violations (Emeakaroha et al., 2012b).

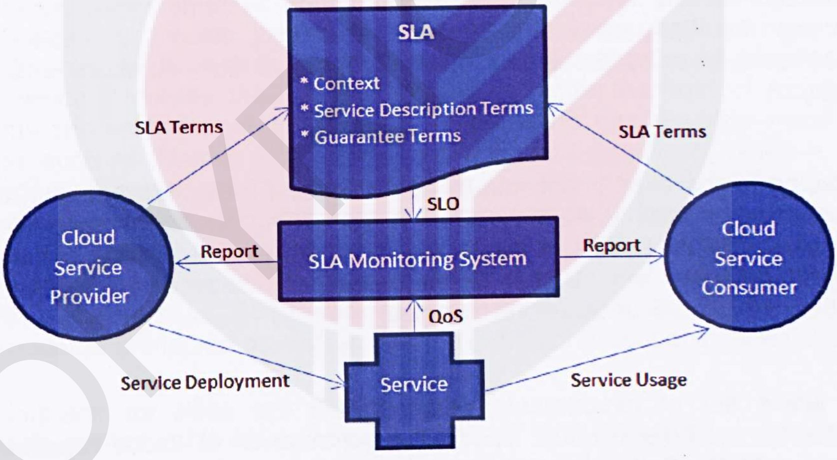

Figure 1.1 SLA in cloud computing

Both service provider and consumer need to monitor the quality of service (QoS) to be sure about SLA validity. Most SLA come with penalty model to compensate the customer in case of violations (Buyya et al., 2010).

'Interval' is defined as a non-overlapping interval of time within the time space of the interaction (Hammadi et al., 2013). There are two types of intervals that should be determined in advance when monitoring the enforcement of SLAs,

(a) *monitoring intervals* between the collection of the monitoring statistics, and (b) *measurement intervals* that aggregate the statistics to measure the service level parameters. High intrusiveness may affect the overall system performances negatively, while low frequency of instrument may cause heavy SLA violations (Maurer et al., 2010). The determination of the intervals is still an open research issue. For example, the monitoring infrastructure of cloud services proposed by Katsaros et al. (2012a) requires the pre-set of the interval of data collection by domain experts. This method heavily depends on the experience of the experts with an insufficiency of rationality and accuracy. Another relevant work is the billing system THEMIS (Park et al., 2013), in which the monitoring report could be presented to consumers when they require the monitoring data. To suggest an optimal measurement interval for detecting applications' SLA objectives violations at runtime, the following two factors are discussed: (a) cost of making measurements; and (b) the benefit of SLA violation detection (Emeakaroha et al., 2012b). The acceptable trade-off between these two factors defines the optimal measurement interval (Emeakaroha et al., 2012b).

The cost value and benefit value of SLA monitoring systems are issues in cloud computing (Emeakaroha et al., 2012b). SLA monitoring systems need resources consumption consisting CPU, Memory, and Storage for execution. The amount of consumed resources by monitoring system is the cost of SLA evaluation. On the other hand, SLA monitoring systems make benefits by detecting the SLA violations because service provider can absorb the infrastructure cost by preventing more numbers of SLA violations and avoid penalty.

According to the experiment results achieved by Emeakaroha et al. (2012b), the number of missed SLA violation is increasing when the measurement interval is increased. This is caused by the fact that the larger measurement interval, the lower the number of measurements made and the higher the number of missed SLA violations detected.

The whole set of experiments demonstrated the need for fine-tuning of monitoring systems to the specific requirements of cloud applications. However, different applications have needs for different measurement intervals, and even though some applications are more stable than other in terms of resource requirements. Therefore, this study aims to optimize the measurement interval of each application at the running time to keep balance between cost and benefit of SLA monitoring systems.

1.2 Problem Statement

Since SLA parameters are usually defined by cloud providers and may comprise of various user-defined attributes, current monitoring infrastructures is lacking of appropriate solutions for SLA evaluation (Emeakaroha et al., 2012b). The first issue is to facilitate mapping the run time data of several metrics by low level tools to numerous SLAs (Emeakaroha et al., 2010a). The existing monitoring infrastructure of cloud services requires the pre- set of the data collection by domain expert. All the data refer to historical data, not at run time data (Katsaros et al., 2012a).The second issue is to determine appropriate monitoring intervals at the application level keeping the balance between the early detection of possible SLA violations and system intrusiveness of the monitoring tools (Emeakaroha et al., 2012b). The third issue is to make a trade-off between the cost of monitoring system and the benefit of SLA violation detection in variable environmental situations (Berral et al., 2010; Emeakaroha et al., 2012b).

Several services are executed in cloud environment at the same time for numerous users. Each service should be monitored at the run-time to evaluate related SLA. Accordingly, SLA monitoring systems face several monitoring targets with different SLAs. Adaptability of SLA monitoring systems is a challenging issue in cloud computing to manage the monitoring threads effectively based on services status. Current monitoring systems have often constant interval value and individual monitoring process for all services and SLAs which decrease the adaptability and increase the cost of monitoring systems (Emeakaroha et al., 2012a). When using the separate monitoring service, such as (Massie et al,. 2004) a single static monitoring policy is used for the entire session that cannot reflect the dynamic changes in the levels of trust and perceived risk. Different services require different measurement interval. (Hammadi et al., 2013; Cuoma et al., 2013; Maurer et al., 2010). This issues has been faced by several studies (Clayman et al., 2011; Katsaros et al., 2011; Kutare et al., 2010; Park et al., 2011; Wang et al., 2011a) in tuning the amount of monitored resources and the monitoring frequency (interval).

The long interval value leads a few numbers of SLA evaluation repetitions; on the other hand, the short interval value causes a lot of repetitions for SLA evaluations (Emeakaroha et al., 2012b). Although the short interval value makes the monitoring loop faster and can detect more number of probable SLA violations, it increases the cost value of SLA monitoring tools. The high interval value has a low overhead cost but it can detect a few numbers of probable SLA violations only. High intrusiveness may affect the overall system performance negatively, while low frequency of instrument may cause heavy SLA violations (Maurer et al., 2010). Therefore the determination of appropriate interval value is a challenging issue because the constant monitoring interval cannot make a balance between cost value and benefit value of SLA monitoring systems (Emeakaroha et al., 2012b).

1.3 Objectives

Following the stated problem, the research objectives of this study are as listed below.

- To design a SLA evaluation framework to increase the adaptability of SLA monitoring system for mapping of several metrics to numerous SLAs.
- To propose a dynamic monitoring interval (DMI) to keep balance between the early detection of possible SLA violations and system intrusiveness of the monitoring tools in variable environmental situations.
- To develop SLA monitoring tool based on proposed DMI to make a trade-off between cost value and benefit value of SLA monitoring system in cloud environment.

1.4 Research Scope

The life cycle of SLA in cloud computing consist of SLA negotiation, SLA evaluation, and SLA expiration. This study is concentrated on SLA evaluation phase. The raw data is collected from resources and the low level metrics are mapped to high level attributes. However, when measured attributes exceed from agreed SLO, the SLA violation is occurred and it should be reported. The main parts of SLA evaluation framework consist of raw-data collecting, low-level metrics to high-level attribute mapping, and run-time evaluating.

The costs and benefits of SLA monitoring system is another scope of this study. Interval value of SLA monitoring system is the central concentrated issue in this area because it has a high impact on cost value and benefit value of monitoring system. SLA monitoring systems currently settled a static interval value to evaluate SLA periodically and detect any probable SLA violations in cloud application. According to Emeakaroha et al. (2012b) different applications have needs for different measurement intervals, therefore, this study will scope its experiment and measurement into finding the dynamic intervals for optimum balance cost and value in SLA evaluations.

1.5 Contributions

The proposed SLA evaluation framework is expected to increase the adaptability of SLA monitoring, and each SLA is evaluated by individual monitoring thread and its own established interval value. The proposed DMI should able to optimize and be adapted as interval value for monitoring threads. As a result, the developed SLA monitoring tools would make a trade-off between cost value and benefit value of SLA monitoring system. The probable SLA violations are detected by developed monitoring tools while it has affordable cost for service providers. The proposed DMI enables the developed monitoring tools to adapt the monitoring interval based on environmental situations at running time and DMI leads to achieving an affordable cost-benefit assessment results as compared to static interval values in variable environmental situations.The main contributions of this study are:

- The proposed SLA monitoring framework to evaluate agreed SLAs in cloud computing at run time.
- The establishment of dynamic monitoring interval (DMI) formula to adapt the interval value of monitoring system for different situations.
- The development of an SLA monitoring tool which is developed based on proposed DMI to monitor the running cloud services.

1.6 Thesis Organization

This thesis is organized in six chapters. The first chapter is the Introduction Chapter which contains the background, problem statement, objective, scope of research and contributions. A comprehensive literature review is presented in Chapter 2. The related works and published papers are selected based on the research scope and objectives to be discussed in Chapter 2. SLA monitoring systems in cloud computing, the monitoring interval, and the cost-benefit of SLA evaluation are the main subjects of Literature Review Chapter. Chapter 3 presents the research methodology consisting research activities, experimental research, experiment design, testbed, and analysis criteria. The proposed SLA monitoring framework, proposed dynamic monitoring interval, and the development of SLA monitoring tools are described in Chapter 4. Chapter 5 shows the results of experiment and discusses about analyzed data. The results of developed monitoring tools in certain environmental situations are presented and they are discussed based on the cost-benefit assessments. Finally, Chapter 6 presents the conclusion and future works.

REFERENCES

Aijaz, F., Shaikh, S. J., and Walke, B. (2010). "A framework for multi-interfaced service level agreements on mobile devices." *Presented at Communication Systems (ICCS), 2010 IEEE International Conference on.*

Alhamad, M., Dillon, T., and Chang, E. (2010a). "Conceptual SLA framework for cloud computing." *Presented at Digital Ecosystems and Technologies (DEST), 2010 4th IEEE International Conference on.*

Alhamad, M., Dillon, T., and Chang, E. (2010b). "SLA-based trust model for cloud computing." *Presented at Network-Based Information Systems (NBiS), 2010 13th International Conference on.*

Alsulaiman, L., and Alturki, R. (2012). "Monitoring multimedia quality of service in public Cloud Service Level Agreements." *Presented at Multimedia Computing and Systems (ICMCS), 2012 International Conference on.*

Andrade, A., and Melo, A. C. M. d. (2011). "A WS-Agreement-based QoS Auditor Negotiation Mechanism for Grids." *Presented at Proceedings of the 2011 IEEE/ACM 12th International Conference on Grid Computing.*

Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Nakata, T., Pruyne, J., Rofrano, J., Tuecke, S., and Xu, M. (2007). "Web services agreement specification (WS-Agreement)." *Presented at Open Grid Forum.*

Anithakumari, S., and Sekaran, K. C. (2014). "Autonomic SLA Management in Cloud Computing Services", *Recent Trends in Computer Networks and Distributed Systems Security.* Springer, pp. 151-159.

Arockiam, L., Monikandan, S., and Parthasarathy, G. (2011). "Cloud computing: a survey." *International Journal of Internet Computing (IJIC), ISSN*(2231-6965), 1.

Balaton, Z., and Gombás, G. (2003). "Resource and job monitoring in the grid", *Euro-Par 2003 Parallel Processing.* Springer, pp. 404-411.

Balaton, Z., Kacsuk, P., and Podhorszki, N. (2001). "Application Monitoring in the Grid with GRM and PROVE", *Computational Science—ICCS 2001.* Springer, pp. 253-262.

Baliś, B., Bubak, M., Funika, W., Szepieniec, T., and Wismüller, R. (2002). "An infrastructure for Grid application monitoring", *Recent Advances in Parallel Virtual Machine and Message Passing Interface.* Springer, pp. 41-49.

Baliś, B., Bubak, M., Funika, W., Szepieniec, T., Wismüller, R., and Radecki, M. (2004). "Monitoring grid applications with grid-enabled OMIS monitor." *Presented at Grid Computing.*

Balis, B., Slota, R., Kitowski, J., and Bubak, M. (2012). "On-Line monitoring of service-level agreements in the grid." *Presented at Euro-Par 2011: Parallel Processing Workshops.*

Bao, D., Xiao, Z., Sun, Y., and Zhao, J. (2010). "A method and framework for quality of cloud services measurement." *Presented at Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on.*

Barbosa, A. C., Sauvé, J., Cirne, W., and Carelli, M. (2006). "Evaluating architectures for independently auditing service level agreements." *Future Generation Computer Systems*, 22(7), 721-731.

Berral, J. L., Goiri, Í., Nou, R., Julià, F., Guitart, J., Gavaldà, R., and Torres, J. (2010). "Towards energy-aware scheduling in data centers using machine learning." *Presented at Proceedings of the 1st International Conference on energy-Efficient Computing and Networking.*

Bertolino, A., Calabrò, A., and De Angelis, G. (2013a). "Adaptive SLA monitoring of service choreographies enacted on the Cloud." *Presented at Maintenance and Evolution of Service-Oriented and Cloud-Based Systems (MESOCA), 2013 IEEE 7th International Symposium on the.*

Bertolino, A., Calabrò, A., and De Angelis, G. (2013b). "A generative approach for the adaptive monitoring of SLA in service choreographies", *Web Engineering.* Springer, pp. 408-415.

Bertolino, A., De Angelis, G., Kellomäki, S., and Polini, A. (2012). "Enhancing service federation trustworthiness through online testing." *Computer*, 45(1), 66-72.

Boley, H., Tabet, S., and Wagner, G. (2001). "Design Rationale for RuleML: A Markup Language for Semantic Web Rules." *Presented at SWWS.*

Boniface, M., Nasser, B., Papay, J., Phillips, S. C., Servin, A., Yang, X., Zlatev, Z., Gogouvitis, S. V., Katsaros, G., and Konstanteli, K. (2010). "Platform-as-a-service architecture for real-time quality of service management in clouds." *Presented at Internet and Web Applications and Services (ICIW), 2010 Fifth International Conference on.*

Boniface, M., Phillips, S. C., Sanchez-Macian, A., and Surridge, M. (2009). "Dynamic service provisioning using GRIA SLAs." *Presented at Service-Oriented Computing-ICSOC 2007 Workshops.*

Borges, H. P., de Souza, J. N., Schulze, B., and Mury, A. R. (2012). "Automatic generation of platforms in cloud computing." *Presented at Network Operations and Management Symposium (NOMS), 2012 IEEE.*

Brandic, I., Music, D., Leitner, P., and Dustdar, S. (2009). "Vieslaf framework: Enabling adaptive and versatile sla-management", *Grid Economics and Business Models.* Springer, pp. 60-73.

Buck, K., Hanf, D. (2010). "Cloud SLA Considerations for the Government Consumer." *System Engineering at MITRE, Cloud Computing Series.*

Buyya, R., Garg, S. K., and Calheiros, R. N. (2011). "SLA-oriented resource provisioning for cloud computing: Challenges, architecture, and solutions." *Presented at Cloud and Service Computing (CSC), 2011 International Conference on.*

Buyya, R., Pandey, S., and Vecchiola, C. (2010). "Market-oriented cloud computing and the cloudbus toolkit." *Large Scale Network-Centric Distributed Systems*, 319-358.

Cardellini, V., Casalicchio, E., Lo Presti, F., and Silvestri, L. (2011). "Sla-aware resource management for application service providers in the cloud." *Presented at Network Cloud Computing and Applications (NCCA), 2011 First International Symposium on.*

Casalicchio, E., and Silvestri, L. (2011). "Architectures for autonomic service management in cloud-based systems." *Presented at Computers and Communications (ISCC), 2011 IEEE Symposium on.*

Casalicchio, E., and Silvestri, L. (2013). "Mechanisms for SLA provisioning in cloud-based service providers." *Computer Networks*, 57(3), 795-810.

Chazalet, A. (2010). "Service level checking in the cloud computing context." *Presented at Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on.*

Cibrán, M. A., and Verheecke, B. (2003). "Modularizing web services management with AOP." *Orientation and Web Services*, 14.

Cicotti, G., Coppolino, L., Cristaldi, R., D'Antonio, S., and Romano, L. (2012). "QoS monitoring in a cloud services environment: the SRT-15 approach." *Presented at Euro-Par 2011: Parallel Processing Workshops.*

Clark, K., Warnier, M., and Brazier, F. M. (2011). *Self-adaptive service monitoring*: Springer.

Clark, K., Warnier, M., and Brazier, F. M. (2013). "Automated non-repudiable cloud resource allocation", *Cloud Computing and Services Science.* Springer, pp. 168-182.

Clayman, S., Clegg, R., Mamatas, L., Pavlou, G., and Galis, A. (2011). "Monitoring, aggregation and filtering for efficient management of virtual networks." *Presented at Proceedings of the 7th International Conference on Network and Services Management.*

Clayman, S., Galis, A., Chapman, C., Toffetti, G., Rodero-Merino, L., Vaquero, L. M., Nagin, K., and Rochwerger, B. (2010). "Monitoring Service Clouds in the Future Internet." *Presented at Future Internet Assembly.*

Comuzzi, M., Kotsokalis, C., Spanoudakis, G., and Yahyapour, R. (2009). "Establishing and monitoring SLAs in complex service based systems." *Presented at Web Services, 2009. ICWS 2009. IEEE International Conference on.*

Coppolino, L., D'Antonio, S., Romano, L., Aisopos, F., and Tserpes, K. (2014). "Effective QoS Monitoring in Large Scale Social Networks", *Intelligent Distributed Computing VII*. Springer, pp. 249-259.

Cuomo, A., Di Modica, G., Distefano, S., Puliafito, A., Rak, M., Tomarchio, O., Venticinque, S., and Villano, U. (2013). "An SLA-based broker for cloud infrastructures." *Journal of grid computing*, 11(1), 1-25.

Curbera, F., Hallam-Baker, P., Hondo, V. M., Nadalin, A., Nataraj Nagaratnam, I., and Sharp, C. (2006). "Web Services Policy Framework (WS-Policy)." *Policy*.

Daniel, D., and Lovesum, S. J. (2011). "A novel approach for scheduling service request in cloud with trust monitor." *Presented at Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), 2011 International Conference on.*

Dastjerdi, A. V., Tabatabaei, S. G. H., and Buyya, R. (2012). "A dependency□aware ontology□based approach for deploying service level agreement monitoring services in Cloud." *Software: Practice and Experience*, 42(4), 501-518.

Davin, J., Case, J., Fedor, M., and Schoffstall, M. (1990). "Simple network management protocol (SNMP)."

Dhingra, M., Lakshmi, J., and Nandy, S. (2012). "Resource usage monitoring in clouds." *Presented at Proceedings of the 2012 ACM/IEEE 13th International Conference on Grid Computing.*

Didona, D., Di Sanzo, P., Palmieri, R., Peluso, S., Quaglia, F., and Romano, P. (2012). "Automated Workload Characterization in Cloud-based Transactional Data Grids." *Presented at 17th IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems (DPDNS).*

Distefano, S., Puliafito, A., Rak, M., Venticinque, S., Villano, U., Cuomo, A., Di Modica, G., and Tomarchio, O. (2011). "Qos management in cloud@ home infrastructures." *Presented at Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2011 International Conference on.*

Dobson, G., and Sanchez-Macian, A. (2006). "Towards unified QoS/SLA ontologies." *Presented at Services Computing Workshops, 2006. SCW'06. IEEE.*

Emeakaroha, V. C., Brandic, I., Maurer, M., and Dustdar, S. (2010a). "Low level metrics to high level SLAs-LoM2HiS framework: Bridging the gap between monitored metrics and SLA parameters in cloud environments." *Presented at High Performance Computing and Simulation (HPCS), 2010 International Conference on.*

Emeakaroha, V. C., Calheiros, R. N., Netto, M. A., Brandic, I., and De Rose, C. A. (2010b). "DeSVi: An architecture for detecting SLA violations in cloud computing infrastructures." *Presented at Proceedings of the 2nd International ICST Conference on Cloud Computing (CloudComp'10).*

Emeakaroha, V. C., Ferreto, T. C., Netto, M. A. S., Brandic, I., and De Rose, C. A. (2012a). "Casvid: Application level monitoring for sla violation detection in clouds." *Presented at Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th Annual.*

Emeakaroha, V. C., Netto, M. A., Calheiros, R. N., Brandic, I., Buyya, R., and De Rose, C. A. (2012b). "Towards autonomic detection of sla violations in cloud infrastructures." *Future Generation Computer Systems*, 28(7), 1017-1029.

Falasi, A. A., Serhani, M. A., and Dssouli, R. (2013). "A Model for Multi-levels SLA Monitoring in Federated Cloud Environment." *Presented at Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC).*

Ferrer, A. J., Hernández, F., Tordsson, J., Elmroth, E., Ali-Eldin, A., Zsigri, C., Sirvent, R., Guitart, J., Badia, R. M., and Djemame, K. (2012). "OPTIMIS: A holistic approach to cloud service provisioning." *Future Generation Computer Systems*, 28(1), 66-77.

Ferreto, T. C., De Rose, C. A., and De Rose, L. (2002). "Rvision: An open and high configurable tool for cluster monitoring." *Presented at Cluster Computing and the Grid, 2002. 2nd IEEE/ACM International Symposium on.*

Ferretti, S., Ghini, V., Panzieri, F., Pellegrini, M., and Turrini, E. (2010). "Qos–aware clouds." *Presented at Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on.*

Foster, H., and Spanoudakis, G. (2011). "Advanced service monitoring configurations with SLA decomposition and selection." *Presented at Proceedings of the 2011 ACM Symposium on Applied Computing.*

Freitas, A. L., Parlavantzas, N., and Pazat, J.-L. (2010). "A QoS assurance framework for distributed infrastructures." *Presented at Proceedings of the 3rd International Workshop on Monitoring, Adaptation and Beyond.*

Frutos, H. M., and Kotsiopoulos, I. (2009). "BREIN: Business Objective Driven Reliable and Intelligent Grids for Real Business." *IBIS*, 8, 39-41.

Fu, W., and Huang, Q. (2006). "GridEye: A service-oriented grid monitoring system with improved forecasting algorithm." *Presented at Grid and Cooperative Computing Workshops, 2006. GCCW'06. Fifth International Conference on.*

Garg, S. K., Versteeg, S., and Buyya, R. (2013). "A framework for ranking of cloud computing services." *Future Generation Computer Systems*, 29(4), 1012-1023.

Goiri, Í., Julià, F., Fitó, J. O., Macías, M., and Guitart, J. (2012). "Supporting CPU-based guarantees in cloud SLAs via resource-level QoS metrics." *Future Generation Computer Systems*, 28(8), 1295-1302.

Hammadi, A., Hussain, O. K., Dillon, T., and Hussain, F. K. (2013). "A framework for SLA management in cloud computing for informed decision making." *Cluster computing*, 16(4), 961-977.

Hammadi, A. M., and Hussain, O. (2012). "A framework for SLA assurance in cloud computing." *Presented at Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conference on.*

Haq, I. U., Schikuta, E., Brandic, I., Paschke, A., and Boley, H. (2010). "Sla validation of service value chains." *Presented at Grid and Cloud Computing, International Conference on.*

Hazzan, O., Lapidot, T., and Ragonis, N. (2011). *Guide to Teaching Computer Science: An Activity-Based Approach*: Springer Science & Business Media.

Hu, J.-Y., Wu, C.-H., Chu, C.-C., Liang, K.-H., Young, H.-C., Hsu, Y.-Y., Hu, C.-H., and Lin, H.-G. (2011). "Constructing a Cloud-centric Service Assurance Platform for Computing as a Service." *Presented at Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2011 International Conference on.*

Katsaros, G., Gallizo, G., Kübert, R., Wang, T., Fitó, J. O., and Espling, D. (2012a). "An integrated monitoring infrastructure for cloud environments", *Cloud Computing and Services Science.* Springer, pp. 149-164.

Katsaros, G., Kousiouris, G., Gogouvitis, S. V., Kyriazis, D., Menychtas, A., and Varvarigou, T. (2012b). "A Self-adaptive hierarchical monitoring mechanism for Clouds." *Journal of Systems and Software*, 85(5), 1029-1041.

Katsaros, G., Kubert, R., and Gallizo, G. (2011). "Building a service-oriented monitoring framework with rest and nagios." *Presented at Services Computing (SCC), 2011 IEEE International Conference on.*

Katsaros, G., Subirats, J., Fitó, J. O., Guitart, J., Gilet, P., and Espling, D. (2013). "A service framework for energy-aware monitoring and VM management in Clouds." *Future Generation Computer Systems*, 29(8), 2077-2091.

Keller, A., and Ludwig, H. (2003). "The WSLA framework: Specifying and monitoring service level agreements for web services." *Journal of Network and Systems Management*, 11(1), 57-81.

Koller, B., and Schubert, L. (2007). "Towards autonomous SLA management using a proxy-like approach." *Multiagent and Grid Systems*, 3(3), 313-325.

Kuchkuda, R. (1988). "An introduction to ray tracing", *Theoretical foundations of computer graphics and CAD.* Springer, pp. 1039-1060.

Kutare, M., Eisenhauer, G., Wang, C., Schwan, K., Talwar, V., and Wolf, M. (2010). "Monalytics: online monitoring and analytics for managing large scale data centers." *Presented at Proceedings of the 7th international conference on Autonomic computing.*

Lamanna, D. D., Skene, J., and Emmerich, W. (2003). "SLAng: a language for service level agreements."

Lee, C. B., and Snavely, A. (2006). "On the user–scheduler dialogue: studies of user-provided runtime estimates and utility functions." *International Journal of High Performance Computing Applications*, 20(4), 495-506.

Lee, K., Paton, N. W., Sakellariou, R., and Fernandes, A. A. (2009). "Utility Driven Adaptive Workflow Execution." *Presented at Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid.*

Lewin, K. (2009). "Federal cloud computing initiative overview." *General Services Administration.*

Lodde, A., Schlechter, A., Bauler, P., and Feltz, F. (2011). "SLA-Driven Resource Provisioning in the Cloud." *Presented at Network Cloud*

Computing and Applications (NCCA), 2011 First International Symposium on.

Ludwig, H., Keller, A., Dan, A., King, R. P., and Franck, R. (2003). "Web service level agreement (WSLA) language specification." *IBM Corporation*, 815-824.

Massie, M. L., Chun, B. N., and Culler, D. E. (2004). "The ganglia distributed monitoring system: design, implementation, and experience." *Parallel Computing*, 30(7), 817-840.

Mathieu, H., and Biennier, F. (2007). "A service level agreement management architecture to improve an information system, business processes administration and evolution." *Schedae Informaticae*, 16, 69-86.

Maurer, M., Brandic, I., Emeakaroha, V. C., and Dustdar, S. (2010). "Towards knowledge management in self-adaptable clouds." *Presented at Services (SERVICES-1), 2010 6th World Congress on.*

Maurer, M., Brandic, I., and Sakellariou, R. (2013). "Adaptive resource configuration for Cloud infrastructure management." *Future Generation Computer Systems*, 29(2), 472-487.

Maurer, M., Breskovic, I., Emeakaroha, V. C., and Brandic, I. (2011). "Revealing the MAPE loop for the autonomic management of cloud infrastructures." *Presented at Computers and Communications (ISCC), 2011 IEEE Symposium on.*

Michlmayr, A., Rosenberg, F., Leitner, P., and Dustdar, S. (2009). "Comprehensive qos monitoring of web services and event-based sla violation detection." *Presented at Proceedings of the 4th international workshop on middleware for service oriented computing.*

Moses, J., Iyer, R., Illikkal, R., Srinivasan, S., and Aisopos, K. (2011). "Shared resource monitoring and throughput optimization in cloud-computing datacenters." *Presented at Parallel & Distributed Processing Symposium (IPDPS), 2011 IEEE International.*

Nie, G., Xueni, E., and Chen, D. (2012). "Research on Service Level Agreement in Cloud Computing", *Advances in Electric and Electronics.* Springer, pp. 39-43.

Niehörster, O., Brinkmann, A., Keller, A., Kleineweber, C., Krüger, J., and Simon, J. (2012). "Cost-aware and slo-fulfilling software as a service." *Journal of grid computing*, 10(3), 553-577.

Papaemmanouil, O. (2012). "Supporting extensible performance SLAs for cloud databases." *Presented at Data Engineering Workshops (ICDEW), 2012 IEEE 28th International Conference on.*

Park, J., Yu, H., Chung, K., and Lee, E. (2011). "Markov chain based monitoring service for fault tolerance in mobile cloud computing." *Presented at Advanced Information Networking and Applications (WAINA), 2011 IEEE Workshops of International Conference on.*

Park, K.-W., Han, J., Chung, J., and Park, K. H. (2013). "THEMIS: A Mutually verifiable billing system for the cloud computing environment." *Services Computing, IEEE Transactions on*, 6(3), 300-313.

Paschke, A. (2005). "Rbsla a declarative rule-based service level agreement language based on ruleml." *Presented at Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on.*

Patel, P., Ranabahu, A. H., and Sheth, A. P. (2009). "Service level agreement in cloud computing."

Racz, P., and Stiller, B. (2009). "Monitoring of SLA compliances for hosted streaming services." *Presented at Integrated Network Management, 2009. IM'09. IFIP/IEEE International Symposium on.*

Raimondi, F., Skene, J., and Emmerich, W. (2008). "Efficient online monitoring of web-service SLAs." *Presented at Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of software engineering.*

Rak, M., Venticinque, S., Máhr, T., Echevarria, G., and Esnal, G. (2011). "Cloud application monitoring: The mOSAIC approach." *Presented at Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on.*

Rao, J., Wei, Y., Gong, J., and Xu, C.-Z. (2013). "Qos guarantees and service differentiation for dynamic cloud applications." *Network and Service Management, IEEE Transactions on*, 10(1), 43-55.

Ravindran, K., Adiththan, A., and Iannelli, M. (2014). "SLA evaluation with on-the-fly measurements of distributed service implementation over clouds." *Presented at Proceedings of the 6th International Workshop on Principles of Engineering Service-Oriented and Cloud Systems.*

Rimal, B. P., Jukan, A., Katsaros, D., and Goeleven, Y. (2011). "Architectural requirements for cloud computing systems: an enterprise cloud approach." *Journal of grid computing*, 9(1), 3-26.

Roman, D., Keller, U., Lausen, H., de Bruijn, J., Lara, R., Stollberg, M., Polleres, A., Feier, C., Bussler, C., and Fensel, D. (2005). "Web service modeling ontology." *Applied ontology*, 1(1), 77-106.

Romano, L., De Mari, D., Jerzak, Z., and Fetzer, C. (2011). "A novel approach to QoS monitoring in the cloud." *Presented at Data Compression, Communications and Processing (CCP), 2011 First International Conference on.*

Sakr, S., and Liu, A. (2012). "SLA-based and consumer-centric dynamic provisioning for cloud databases." *Presented at Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on.*

Schmidt, S., Steele, R., Dillon, T., and Chang, E. (2005). "Building a fuzzy trust network in unsupervised multi-agent environments." *Presented at On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops.*

Schulz, F., Michalk, W., Hedwig, M., McCallister, M., Momm, C., Caton, S., Haas, C., Rolli, D., and Tavas, M. (2012). "Service Level Management for Service Value Networks." *Presented at Computer Software and Applications Conference Workshops (COMPSACW), 2012 IEEE 36th Annual.*

Shao, J., and Wang, Q. (2011). "A performance guarantee approach for cloud applications based on monitoring." *Presented at Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual.*

Słota, R., Nikolow, D., Młocek, P., and Kitowski, J. (2012). "Semantic-based SLA monitoring of storage resources", *Parallel Processing and Applied Mathematics*. Springer, pp. 232-241.

Stiller, B. (2005). "A generic model and architecture for automated auditing", *Ambient Networks*. Springer, pp. 121-132.

Sun, L., Singh, J., and Hussain, O. K. (2012). "Service level agreement (SLA) assurance for cloud services: a survey from a transactional risk perspective." *Presented at Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia.*

Toueir, A., Broish, J., and Sibilla, M. (2014). "Reconfiguration Patterns for Goal_Oriented Monitoring Adaptation. ", *The Sixth International Conferences on Pervasive Patterns and Applications*, pp. 22-27.

Tosic, V., Patel, K., and Pagurek, B. (2002). "Wsol—web service offerings language", *Web Services, E-Business, and the Semantic Web*. Springer, pp. 57-67.

Tserpes, K., Kyriazis, D., Menychtas, A., and Varvarigou, T. (2008). "A novel mechanism for provisioning of high-level quality of service information in grid environments." *European Journal of Operational Research*, 191(3), 1113-1131.

Vrcek, N., and Brumec, S. (2013). "Role of utilization rate on cloud computing cost effectivness analysis." *Presented at Information Society (i-Society), 2013 International Conference on.*

Walker, E. (2009). "The real cost of a CPU hour"*ImPAC, CPU LEASING*. City: IEEE Computer, pp. 35-41.

Wang, C., Schwan, K., Talwar, V., Eisenhauer, G., Hu, L., and Wolf, M. (2011a). "A flexible architecture integrating monitoring and analytics for managing large-scale data centers." *Presented at Proceedings of the 8th ACM international conference on Autonomic computing.*

Wang, X., Du, Z., Liu, X., Xie, H., and Jia, X. (2010). "An adaptive QoS management framework for VoD cloud service centers." *Presented at Computer Application and System Modeling (ICCASM), 2010 International Conference on.*

Wang, Z., Tang, X., and Luo, X. (2011b). "Policy-based SLA-aware cloud service provision framework." *Presented at Semantics Knowledge and Grid (SKG), 2011 Seventh International Conference on.*

Wieder, P., Butler, J. M., Theilmann, W., and Yahyapour, R. (2011). *Service level agreements for cloud computing*: Springer.

Wood, T., Shenoy, P., Venkataramani, A., and Yousif, M. (2009). "Sandpiper: Black-box and gray-box resource management for virtual machines." *Computer Networks*, 53(17), 2923-2938.

Wood, T., Shenoy, P. J., Venkataramani, A., and Yousif, M. S. (2007). "Black-box and Gray-box Strategies for Virtual Machine Migration." *Presented at NSDI.*

Wu, L., and Buyya, R. (2010). "Service level agreement (SLA) in utility computing systems." *Performance and Dependability in Service Computing: Concepts, Techniques and Research Directions*, 1-25.

Ye, L., Zhang, H., Shi, J., and Du, X. (2012). "Verifying cloud service level agreement." *Presented at Global Communications Conference (GLOBECOM), 2012 IEEE.*

Yeo, C. S., and Buyya, R. (2007). "Pricing for utility-driven resource management and allocation in clusters." *International Journal of High Performance Computing Applications*, 21(4), 405-418.