INT J COMPUT COMMUN, ISSN 1841-9836
8(1):42-49, February, 2013.

Datastores in Cloud Governance

A. Copie, T.-F. Fortiş, V.I. Munteanu

Adrian Copie, Teodor-Florin Fortiş,
Victor Ion Munteanu
1. West University of Timişoara
Romania, Timişoara, bvd. V.Pârvan 4, and
2. Institute e-Austria, Timişoara
Romania, Timişoara, bvd. V.Pârvan 4
adrian.copie@info.uvt.ro,fortis@info.uvt.ro
vmunteanu@info.uvt.ro

Abstract:
The Small and Medium Enterprises benefit now, due to the scale adoption of Cloud Computing, from an emerging market where they can associate and collaborate to form virtual enterprises or virtual clusters, aiming to compete with the large enterprises and provide tailored IT solutions for their customers. However the lack of standardization for the cloud services and technologies leads to a myriad of different components that cannot be easily set to work together in the absence of a real cloud governance solution. Cloud governance acts like a catalyst to allow Small and Medium Enterprises to easily manage and optimize their services infrastructure, and to facilitate collaboration in a clustered or virtual-enterprise environment.
We have proposed a Cloud Governance architecture based on mOSAIC's multi-agent Cloud Management solution. The Cloud Governance solution relies on various datastores that are responsible with maintaining and managing a set of crucial data that are used during the cloud governance process . Our paper is focused to analyze and emphasize the requirements that must be fulfilled by different database systems in order to have a reliable storage system and also to suggest a concrete solution.
Keywords: Cloud Computing, Cloud Governance, Datastores, Databases

1 Introduction

In the last years, Cloud Computing becomes a common paradigm, due to the straightforward way of resource provisioning, dynamically scalability, service orientation and its simple pay-as-you-go financial model. Along with these characteristics, in [1] are identified others like strong fault tolerance, loosely coupling, virtualization, ease of use and the link with the business model which lead to the development of new business models [2], [3].

The selected architectural model allows SMEs to collaborate and associate in virtual enterprises or virtual clusters and expose services in direct competition with the large enterprises. However, there are still issues to overcome in order to benefit of this kind of collaboration due to the diversity of services offered by the cloud providers and the lack of standardization.

Efforts have been made in this direction, different solutions that aim to abstract the characteristics of some existing cloud providers have been released: mOSAIC[1], CloudFoundry[2], Morfeo 4CaaSt[3], ActiveState Stackato[4], OpenShift[5], Reservoir[6], SLA@SOI[7] and many others . Even if they are a step forward in the process of the cloud services standardization, [4], [5] and [6] suggest

[1] http://www.mosaic-cloud.eu
[2] http://cloudfoundry.org
[3] http://4caast.morfeo-project.org
[4] http://www.activestate.com/stackato
[5] http://openshift.redhat.com/app/
[6] http://www.reservoir-fp7.eu/
[7] http://sla-at-soi.eu/

complementary services related to the Cloud Management which assures that the resources in the cloud are used optimally and properly interact with the users and other services. Taking into account the growth rate of the cloud services, [7], [8], [9] reveal the need a better integration and a demand for complementary mechanism for Cloud Management: *Cloud Governance*. This is an evolutionary step in the Service Oriented Architecture (SOA) which makes possible the awareness of the existence of the cloud services for the potential consumers. Due to its similarities, the awareness mechanism rely on different specialized datastores which hold critical information and which have various requirements in order to correctly function and offer reliable data. However, according to [11] there must be a clear separation between the management and governance processes, because they describe different activities, have distinct goals and involves different organizational structures.

2 The mOSAIC Project

'Open source API and Platform for multiple Clouds' (mOSAIC) is a FP7-ICT project that aims to provide an open-source platform and an API which abstracts the particularities of various cloud providers and encourages the applications development based on the cloud-programming paradigm. Its main components are the *Cloud Agency* which is an embedded Cloud Management solution based on agents that is designed to negotiate cloud resources and provide them to the second component called *mOSAIC platform* which consumes them by offering a cloud-oriented application development framework.

3 Cloud Governance

Our proposed Cloud Governance architecture is based on variuos proposals, including Distributed Management Task Force's (DMTF) white papers [6], [10] and is built around the mOSAIC's Cloud Agency. The Cloud Agency is a core Cloud Management component in the mOSAIC platform that exposes its functionality as a service that is consumed inside our Cloud Governance component (Figure 1).

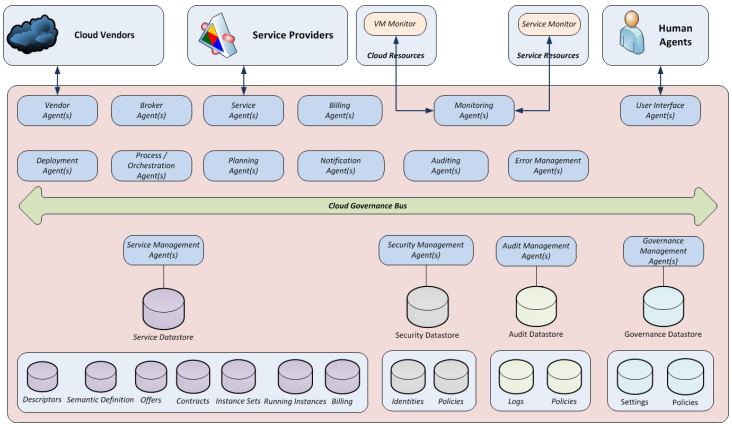

Figure 1: Cloud Governance

The architecture reveals several distinct components like the Cloud Management solution based on mOSAIC's Cloud Agency, the Cloud Governance Bus, the Cloud Governance functional modules and the datastores used to persist various data during the governance process. The Cloud Agency's main role is to assure the management of the resources but also to perform

Table 1: General Requirements for the Datastores in Cloud Governance

Crt.No.	Requirement
S1	Cost optimisation by trying to find solutions in order to minimize the costs with the storage and virtual machines in the cloud through intense multiplexing
S2	High performance in terms of throughput, small latency, scalability independent of the data size and the dynamic of the workload
S3	Security (confidentiality, integrity, privacy)
S4	High availability
D1	Simple internal API in order to expose a relative small number of methods to be used by the governance agents

various SLAs monitoring activities. The messages and data exchange between the cloud services is made through the Cloud Governance Bus. However the governance process is realized using four specialized agencies namely *Service Management*, *Security Management*, *Audit Management* and *Governance Management*. The Service Management takes care about the lifecycle of the services registered in the cloud governance environment. Security Management handles the identities of the service consumers and provides the security tokens used to access the requested services. The Audit Management agency processes the audit information obtained from the cloud management component and in the same time allows the access to a wide range of logs generated by the interacting components. Finally, the Governance Management coordinates the entire governance activity based on a set of rules, policies and settings.

4 Datastores in Cloud Governance

Every cloud governance agency handles specific data, being in direct relation with a dedicated datastore called namely *Service Datastore*, *Security Datastore*, *Audit Datastore* and *Governance Datastore*.

Every datastore is in charge with keeping crucial information related to the functionality of the entire system but also very sensitive from the privacy and confidentiality point of view, such as credentials, contracts, partners, policies, etc. In the same time the data storage system must be very responsive in what concerns the time took to perform various operations over the data set and also must offer a good performance in terms of bandwidth, to accommodate with the processed data flow. This is why the data storage systems must fulfil some general requirements, as revealed in Table 1.

The general requirements for the data storage systems have been divided in two categories: requirements related directly to the intrinsic characteristics of the storage like cost optimization, high performance, high availability, high security and also requirements related to the development phase, like a very simple and intuitive API. These requirements apply to all the datastores inside the governance environment and, from case to case, there are other specific requirements that will be exposed in the appropriate section.

4.1 Governance Datastore

The cloud governance implies the use of various policies related to the way in which the cloud services interact with the consumers, together with different types of constraints that limit the

access to the underlying resources, or even functionality of the governance process building blocks. There are several types of policies that are administered by the Cloud Governance component like Service Level Agreements (SLAs) and Service Level Objectives (SLOs) and which represent the cornerstone of the cloud governance process [4], [6]. The relevant constraints involved in the decision taking process are deployment constraints, data residency constraints, auditability constraints and security constraints. After two partners, namely the provider and the consumer, agree about a set of constraints, they will further govern the interaction between the entities involved in the contract.

The Governance Management Agency coordinates the entire activity performed by the Cloud Governance component based on different policies, constraints and system settings that are stored in the Governance Datastore like access policies, virtual machine settings, security credentials, interaction policies, etc.

Every cloud service provider has its own policies and constraints that will be published and used during the service lifecycle. Some of the policies and constraints could be editable by the service consumers in order to customize them, like access control lists, some of them are not, like Quality of Services, all of them being part of the request and offer operations. The information represented by the policies and constraints is usually structured, being related to the guaranteed functionally parameters of the services, security policies, configuration parameters of the virtual machines on which the agents are executed and many more.

The data persisted in the Governance Datastore is critical to the whole governance process, therefore requirements of the storage system are also in terms of highly availability and reliability. Some of the data is very sensitive so a level of encryption must be added. In the same time, this datastore must maintain relations with other storage systems which recommends a flexible graph database to hold the information related to governance. This choice will assure high performances in terms of records processing, allowing to be installed either in the private cloud on commodity servers or in the public cloud through virtual machines.

4.2 Services Datastore

This is the most complex datastore, part of the cloud governance process. It could be seen like a service itself offering essential interfaces for the cloud services that want to register in the Cloud Governance component and then to be discovered and used, together with other additional features related to offers, contracts, audits and billing. Figure 2 depicts the schematic of the Service Datastores.

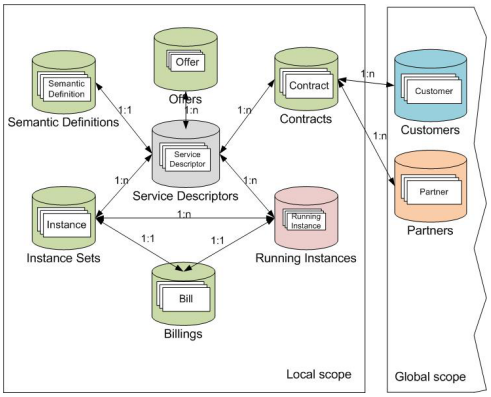

Figure 2: Services Datastore and their internal relationships

The Service Datastores, through their specific mechanisms and information that compound

them, facilitates the services publishing and discovery processes. Like in the case of SOA Governance, Cloud Governance relies on a Service Repository model in which the services that want to offer their functionality must register and must offer methods to be discovered by the consumers. This datastore is also tied with global catalogues that maintain information about the system's customers and the partners. The character of the information contained in the Service Datastore is extremely heterogeneous, holding many kinds of specific information, grouped in several sub-components:

- Service Descriptors. The services functional and non functional requirements are described through service descriptors and they are persisted inside the Descriptors component. This is usually a structured information, the service descriptors contains basically information of the same type and must be highly query-able in order to offer flexible information about the contained services.
- Semantic Definitions. Besides the syntactic description stored inside the Descriptors component, the cloud services are described and defined from their semantic point of view through semantic descriptors and stored in the Semantic Definition sub-component.
- Offers. The relation between provider and consumer usually benefit from a resource model that describes what the service can offer in terms of functionality. Service templates are used to describe in a generic form what a provider can offer, but when the template contains information about a specific provider, the template becomes an offer and it is persisted for future use in the Offers sub-component. Because the information is based on generic templates, it could be structured so an appropriate storage system could be used.
- Contracts. When a consumer looks for a specific service and discover it in the Services Datastore, it consults its offer, agrees with the SLAs and if it decides to use that services the offer becomes a contract that is also stored inside the Contracts sub-component. For a specific service, more than one contract could exist.
- Instance Sets. After a cloud service is contracted, in order to be effectively used it is instantiated and an additional information called Service Instance is generated. This information is an aggregate of trading, billing and deploying information which is stored in a separate sub-component of Services Datastore called Instance Sets.
- Running Instances. The information about the running instances of the cloud services is very much the same as the one contained in the Service Instances and is persisted in the Running Instances sub-component
- Billing. For every contract, billing information is generated in accordance with the parameters agreed during the contracting phase and this information is located in the Billing sub-component.

Taking into account the format of the data persisted in the individual database components in the Service Datastores, together with the general requirements already exposed, the choices in terms of database types are

- NewSQL database for the Service Descriptors database
- RDF datastore for the Semantic Definitions database in order to benefit from the specialized SPARQL query language
- Graph Database for the Offers, Contacts, Instance Sets, Running Instances and Billing databases

4.3 Security Datastore

The Cloud Governance involves distinct entities, components, agents running in behalf of various parties, therefore the security is one of the most important concerns in order to assure a safe and trustful environment. Inside the Cloud Governance system we distinguish two security planes: the security of the cloud services, usually assured by the service providers and the intrinsic security of the Cloud Governance component.

The access to the cloud services must be done in a secure way, only authorized consumers having the rights to use the provided functionalities, sometimes in a granular way through various access rights and security policies. The Security Management Agent is responsible for the credentials management for the authentication and authorization process together with providing the security tokens used to access different cloud resources after the successful authentication process. All the credentials and security policies governing the cloud services in the system are stored in the Security Datastore.

In most cases, the services are offered on a pay-as-you-go schema, being accessed by providing credentials in the form of user name and password, or in terms of security keys. This special kind of data is kept inside the Security Datastore. To add an extra security layer over the persisted data, the information contained in the Identities sub-component must be encrypted and the passwords hashed. The security policies used to consume the cloud services are related to the consumer identities and establish what kind of actions are allowed over a given resource by an authenticated consumer.

The appropriate database type to hold the security information is NewSQL, allowing to be implemented either in the private or public cloud.

4.4 Audit Datastore

The services involved into a complex cloud system that requires governance are provided on a paid scheme, and must respect an SLA previously negotiated and agreed between the cloud service provider and the consumer. Usually the service is monitored only on the provider side, such that the client is not aware when the cloud service does not respect the parameters that are stipulated in the contract.

Through the accepted SLA, different metrics are established in order to facilitate the service control and monitoring mechanisms. The mOSAIC's Cloud Agency component provides the required functionality that allows the service monitoring on the consumer side and to permanently compare the measured parameters with the contractual ones that is leveraged by the Audit Management Agent which obtains all the necessary information about a specific service. Along with the performance measurements, billing information is generated based on the time spent for a cloud service to perform specific tasks together with audit information , triggered when the cloud service is accessed by various entities. The audit mechanism is working based on various policies established at the service level. All these data together with the audit policies is stored in the Audit Datastore component.

Usually the data contained in the Audit Datastore is not structured, but has a large volume depending on the number of the performance parameters monitored and the sampling frequency. Because many decisions in the process of governance are taken based on the data analysis, it is recommended that the storage system that host this data to offer high availability, high writability and scaling, which is better achieved by a key-value datastore.

5 Conclusions and Future Works

Cloud Computing adoption embraced by the Small and Medium Enterprises makes possible the access on new markets where they can associate in virtual clusters or virtual enterprises in order to be able to compete with the large enterprises, by offering complex solutions, tailored on the customers specific needs. We have proposed a Cloud Governance architecture based on mOSAIC's Cloud Management components that aims to manage and govern the services infrastructure. This government solution relies on a set of datastores that maintains and manage various data produced and consumed during the services lifecycle.

Our paper focused in expanding the information contained in the most important datastores in Cloud Governance, determining the requirements for every database that form the datastores and suggests the appropriate database types.

This paper is a base for a future work consisting in the concrete implementation of a Cloud Governance component.

Acknowledgments

This work was partially supported by the grant of the European Commission FP7-ICT-2009-5-256910 (mOSAIC), The views expressed in this paper do not necessarily reflect those of the corresponding projects consortium members.

Bibliography

[1] C. Gong, J. Liu, Q. Zhang, H. Chen, and Z. Gong, The characteristics of cloud computing, in W.-C. Lee and X. Yuan (Eds.), ICPP Workshops, IEEE Computer Society, pp. 275-279, 2010.

[2] C. Weinhardt, A. Anandasivam, B. Blau, N. Borissov, T. Meinl, W. Michalk, J. Stoer, Cloud Computing A Classification, Business Models, and Research Directions, *Business and Information Systems Engineering*,ISSN: 1867-0202, 1(5):391-399, 2009.

[3] C. Weinhardt, A. Anandasivam, B. Blau, and J. Stoer, Business Models in the Service World, IEEE IT Professional, Special Issue on Cloud Computing, ISSN: 1520-9202, 11(2):28-33, 2009. [Online].
Available: http://dx.doi.org/10.1109/MITP.2009.21

[4] Cloud Computing Use Cases Group. (2010, July) Cloud computing use cases white paper. [Online]. Available: http://opencloudmanifesto.org/Cloud Computing Use Cases Whitepaper-4 0.pdf

[5] Use Cases and Interactions for Managing Clouds. Distributed Management Task Force, (2010, June) [Online]. Available: http://www.dmtf.org/sites/default/files/standards/documents/DSP-IS0103 1.0.0.pdf

[6] DMTF. (2010, June) Architecture for Managing Clouds. Distributed Management Task Force. [Online]. Available: http://dmtf.org/sites/default/files/standards/documents/DSP-IS0102 1.0.0.pdf

[7] P. Wainewright. (2011, August) Time to think about cloud governance. [Online]. Available: http://www.zdnet.com/blog/saas/time-to-think-about-cloud-governance/1376

[8] S. Bennett, T. Erl, C. Gee, R. Laird, A. T. Manes, R. Schneider, L. Shuster, A. Tost, and C. Venable, SOA Governance: Governing Shared Services On-Premise in the Cloud. Prentice Hall/PearsonPTR, 2011. [Online]. Available: http://www.soabooks.com/governance/

[9] T. Cecere. (2011, November) Five steps to creating a governance framework for cloud security. Cloud Computing Journal. [Online]. Available: http://cloudcomputing.sys-con.com/node/2073041

[10] DMTF. (2010, June) Use Cases and Interactions for Managing Clouds. Distributed Management Task Force. [Online]. Available: http://www.dmtf.org/sites/default/files/standards/documents/DSP-IS0103 1.0.0.pdf

[11] ISACA. (2012, March) COBIT 5 - Introduction. [Online]. Available: http://www.isaca.org/COBIT/Documents/COBIT5-Introduction.ppt