Cloud Computing Brokering Service: A Trust Framework

Service Level Agreements: An Analytical Study in Progress

Prashant Khanna
Institute of Engineering and Technology
JK Lakshmipat University
Jaipur, India
e-mail: perukhan@gmail.com

Budida Varahala Babu
Institute of Engineering and Technology
JK Lakshmipat University
Jaipur, India
e-mail: profbvbabu@gmail.com

Abstract— The paper highlights existing research voids in defining and designing binding and enforceable service level agreements (SLA) between three actors in the cloud computing framework defined by NIST – the cloud brokers, the cloud consumers and the cloud providers. The paper presents a techno-managerial perspective to the issue of how cloud brokers would handle service provisioning and whether binding service level agreements would be useful tools for the NIST cloud framework to function. A template constituent framework is also recommended as part of this ongoing study.

Keywords—Cloud Brokers; Cloud Computing; SLA; Service Provisioning; Trust.

I. INTRODUCTION

Cloud Computing is an emerging computing paradigm that promises to change the landscape of the present service models on offer in provisioning of Information Technology services. The “Cloud”, as a term has found prominence in an increasingly large number of publications, both in the academia as well as in industry literature. It is a buzz word and the buzz is getting louder by the day. The definitions of cloud computing are many, and varied. The industry has, only in late 2011, finally decided to accept one that was proposed by National Institute of Standards and Technology (NIST), U.S. Department of Commerce [1]. As per the Draft Computing Technology Roadmap published by NIST, Cloud Computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. The definition has listed five essential characteristics that would be common to all cloud computing services, namely: on-demand self service, broad network access, resource pooling, rapid elasticity and measured service. It recommends three service models: Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS), and four deployment models i.e. private, public, community and hybrid clouds.

The reference architecture in the NIST document highlights interactions amongst these entities and provides a companion cloud computing taxonomy detailing the definitions and relationships of a control vocabulary. The document also identifies five major actors to enable the reference model to work, namely, cloud consumer, cloud provider, cloud carrier, cloud auditor, and cloud broker. Each actor is an entity (a person or an organization) that participates in a transaction or process or performs tasks in cloud computing. A lot has been said and written about the model and the way the players interact in this model to derive services. Each of the players have been defined and redefined in literature and the use case(s) to make the model successful has also been commented upon extensively. Amongst the actors defined in the NIST model [1], the cloud broker was an add-on after much thought. Gartner, in a report in 2011 [2] indicated that cloud brokering services in the cloud service marketplace is emerging as a promising low-risk business model for offering new and value-added services through cross provider service delivery and partnership. This assertion has made a major impact on the industry as well as the academia.

Any service brokering architecture, in general, must have the ability to support a service delivery infrastructure for integration, delivery and management of composite services in a multi-provider heterogeneous networks environment. It is no different in the cloud service provisioning environment. In the present stage of evolution of the cloud as a repository of services, this provisioning is far from being ideally achieved. The cloud paradigm is currently in a state of transition and multiple players are trying to dominate the service delivery scene. The cloud providers are competing with the cloud brokers to deliver the intended service to the cloud consumer, but this model of business to consumer interaction is not bearing the desired results due to multiple barriers of scale and other managerial issues. This research on the subject, supported by the industry reports indicate that the player who is likely to emerge as the principal stake holder in provisioning and arbitraging of services as a truly elastic and dynamic package for the consumer would be the cloud broker. Such service provisioning is already appealing to the small and medium business entrants who are not yet as big as Google or Amazon, but have the understanding of how the cloud works [2]. Forrester [3], in their annual report in 2011, also cite brokering services in the cloud to be the next game changer in the service provisioning space. However, the present state of cloud implementation is highly proprietary and private, akin to islands of highly autonomous island solutions which do not have any linking ferry services which can carry the inhabitants across. The cloud brokering
service available today is thus confined to a miniscule subset of matching services that are seamlessly able to speak to each other. There is a serious void in interoperability between cloud solutions that are not been addressed by the present generation of brokering service providers, either due to technological incompatibilities or due to managerial issues. The present NIST framework for the cloud-based service model, as others similar frameworks, are based on adopting managerial practices in organizations which are implemented by using a preferred underlying technology. This is truer today with the inclusion of the cloud broker as an actor in the models under consideration.

We appreciate that this is as much a managerial issue as it is a technical one. This paper and research is an attempt to highlight existing research voids and present a technomanagerial perspective to the issue of how cloud brokers would handle service provisioning and whether binding service level agreements (SLA) would be useful tools for the NIST cloud framework to function. This is a work in progress and it is anticipated that the research would result in proposing a framework that would make service clouds talk to each other under a universally acceptable interoperability standard, where enforceable and automated SLA become corner stones for provisioning of dynamic and elastic services amongst the actors enumerated in the NIST model.

The structure of the remaining paper will be as given below: Section II presents the Literature Survey on the topic of SLA in the cloud and its relevance to the cloud brokering services. Section III addresses the constituents of an SLA within a service oriented business model. Section IV provides the details of a framework in making for an enforceable SLAs amongst cloud brokers, cloud consumers and cloud providers for efficient service provisioning. Sections V and VI conclude current findings and highlights future work directions.

II. LITERATURE SURVEY

Though Cloud computing is a highly studied topic today and a large body of research has gone into studying specific standards of interoperability amongst clouds and how they are to be achieved, the aspects of brokering services to the end client from amongst those available is finding refereed status only recently [4]. A cloud broker has been described as an entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers.

Existing work in literature primarily stress on using SLAs to guarantee consumer of cloud services a level of performance, that is defined by abstract metrics, directly from the cloud service providers to the end client or cloud consumers [9], [10], [11]. There is an apparent void in research on SLA formulation strategies between the cloud service broker and the cloud consumer and between the cloud service broker and the cloud service provider. This research is an attempt to highlight the research void and recommend a framework which can be developed for creation of enforceable and implementable SLAs in the cloud paradigm.

The architecture of the cloud, whether public, private, community or hybrid, would make it non trivial to propose and implement a framework for creating of such binding frameworks in the absence of accurate measuring and monitoring mechanisms for provision of services. This is especially true for a use case when the broker is aggregating and arbitraging services from multiple cloud service providers and packaging them as a service bundle for the end client. Previous work on the subject include [5], [6] and [7] that pertains to SLA formulation, but does not address the aspect of the cloud brokering actor’s role in the provisioning of services. Alhamad [9] [10] discusses the aspect of SLA and performance measurement in his recent findings but does not address the issue in the perspective of how a broker would become a party to the SLA agreement between the end user or the cloud consumer and the cloud service provider. In [25], Alhamad describes a conceptual framework for SLA in the cloud computing paradigm, but the same is silent on the aspect pertaining to Brokers in the service model. Other work on SLA management and creation includes [11], which describes an approach for negotiating and creating SLA between infrastructure providers and service providers. In [12], Parrilli provides a legal perspective on the aspect of SLA provisioning in the European Union and how the rules on jurisdiction provided by the Regulation 44/2001 where two general distinctions are drawn in order to determine which (European) courts are competent to adjudicate disputes arising out of a SLA. The former is between Business to Business and Business to Consumer transactions, while the latter is in regard to contracts which provide a jurisdiction clause and contracts which do not.

A recent work by Wang et al. [13] addresses the aspects of multi-variable SLA based metrics that manages resource scheduling for application provisioning on the cloud. They also recommend a reputation based system for selecting a cloud provider. In [14], Salvatore et al. discuss a framework for broker assigned SLA management service with a novel high level abstraction model has been recommended. They recommend an architectural design for a system named Cloud Agency that aims to respond to the need for Resources management and offers added value to the existing Cloud services. The proposed system is in charge of brokering the collection of Cloud resources from different providers that fulfills requirements of user's applications as a best effort service. The user is able to delegate to the Agency the necessary checks of the agreement fulfillment, the monitoring of resource utilization and eventually necessary re-negotiations. In [15], Balakrishnan and Somasundaram propose a broker framework where SLA enabled broker evaluates the number of resources available in the environment and the number of policies per resource that need to be implemented. The results presented in the paper indicate that the inclusion of SLA affects the resource selection behavior of the broker. The paper is however silent on the methods to control the affect using an SLA. It does however indicate that the overall performance of the system improves in terms of job throughput with an extra overhead in request processing due to the presence of a broker. These
results are shown on a grid sharing environment and major differences exist in the business model used for the grid service provisioning and cloud service provisioning model.

A number of publications, post 2010 [8], [9], [13], [23], [24], [25] are either addressing the aspect of SLA management for brokering services at the level of a resource scheduler, or abstractions of the same when lifted from the grid computing era. The industry is viewing SLA performance management and service provisioning as a combination of availability parameters and associated factors. The carry forward of concepts of web service based SLAs in literature is also evident while drafting cloud based SLAs in recent papers. However, this research on the topic indicates that the business model of provisioning of these two frameworks is very different and mapping the two under the same head would be a mistake. The same has been asserted by NIST [1]. Quantifiable system level metrics like QoS, CPU utilization, assured storage space, scale up and scale down time in terms of elasticity of service, besides some metrics of security also find mention in industry white papers when they refer to enforceable SLAs. Recent literature also highlight the abstract and non quantifiable aspects of performance management and binding of service issues by cloud service brokers while terming the environment of cloud computing turbulent [16].

The purpose of this paper is to highlight the research gap existing in SLA formulation between the cloud broker-consumers and broker-provider combine. It presents the research done thus far and the likely line of further research to address the void. The researchers believe that the solutions to finding or evolving a framework for enabling such enforceable SLAs would be a combination of adopting appropriate managerial practices by the consumers and incorporating the best available technological means for monitoring and measuring the services available in the cloud. This paper thus presents a techno-managerial perspective to the issue.

The perspective adopted in this paper is that of a cloud broker. It is directed towards a cloud consumer and a cloud provider, when seen from a cloud broker’s angle. This paper does not discuss the implication of a binding SLA between the cloud broker and cloud auditors or the cloud carriers. The relationship impact on these actors from the perspective of the broker will be done as a separate study in future.

III. SLA WITHIN A SERVICE ORIENTATION MODEL

An agreement is always based on a measure of trust. Trust concepts have been defined differently when used in varying contexts. Economists, lawyers and information technologists tend to view trust in different light. Numerous models are proposed in literature that attempt to solve the problems that arise when two parties need to establish a business relationship between them. Hussain and Chang [17] highlight the confusion in literature around the concept of trust. The acceptable definition of trust in a common usage scenario is succinctly provided by Dasgupta in [18] where he defines trust as “the expectation of one person about the actions of others that affects the first person’s choice, when an action must be taken before the actions of others are known.” This paper considers the interaction between the cloud actors in the same context. Gambetta [19], on the other hand, states that “trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action.” In the cloud paradigm this relationship maps to the level of trust that exists between the actors involved in the services provisioning.

A. Trust as a Base for Enforcable SLAs

SLAs are based in an inherent trust relationship. SLAs are legal and formal documents which presents the manner in which a relationship between two entities would evolve and be conducted during normal and extreme circumstances. Service providers use SLA as a foundation to optimize the use of resources available at their disposal, while ensuring that the necessary levels of service, as defined in the SLA is delivered to the consumer. The cloud consumer on the other hand uses a SLA to assure themselves of a minimum level of service, which gets enumerated in the SLA that defines the relationship. For the service industry, SLAs must be modeled around a series of related metrics which govern performance in the specific industry. More specifically, an SLA must clearly define components that govern the relationship between the players. An SLA format should illustrate the following:

- Describe a service in unambiguous terms so all stakeholders understand the implication and expectations from the service.
- Present the level of performance of service in terms of metrics.
- Define a monitoring mechanism that would monitor and report if the defined service levels are being provisioned and available to the consumers.
- A mechanism for measurement of the services being provisioned. It is essential that the process is acceptable to all the players involved in the process.
- Provide a framework for imposing penalty due to diversions from the stated terms in the SLA.
- Provide a mechanism that allows the parties engaged in the SLA to interact and meet on common ground in the event of a dispute.
- Duration of implementation and validity of the SLA.

B. SLAs in the Service Industry Framework

The researchers believe that that for the service industry, and especially for the cloud based business model, an SLA must define adherence to some other common metrics. These metrics need customization based on the kind of services needed by the broker (arbitraged or intermediated). Some of the metrics, which can be included in formulating an effective and enforceable SLA, are presented below:

- Response levels in terms of time for service provisioning.
- Cost of provision of service to the end user.
• Service problem reporting and hierarchy of ticket resolution.
• Resolution mechanisms.
• Monitoring and service reporting accountability in terms of resources responsible for the monitoring and adherence within the time frame agreed upon.
• Liabilities of the service provider in case the desired services are not delivered.
• Terms and taxonomy that is agreed upon by the consumer of services and the other actors involved in the process.
• Conditions extraneous to the agreement which have a binding bearing on the SLA.

C. Factors that Fail SLA based Relationships

It is also pertinent to appreciate factors that have been found to be primary reasons for SLA based relationships to fail at times. Industry literature indicates non-optimal business deals that fall through, do so due to ill conceived or poorly researched SLAs [3]. This research deduced that issues common in such failed SLAs based relationships include:

• Ambiguity in differentiating between results and efforts by the service provider.
• Unclear and incomplete service specifications in the SLA lead to dissimilar level of understanding between the service provider and service consumers and other actors involved in the process.
• Incorrect people in the hierarchy creating and approving the SLA.
• Lack of agreement on common taxonomy and terms of reference.
• Lack of trust after a service related issue between the cloud consumer and cloud provider.

Dinesh [20] cites the three different approaches or models used to create a binding SLA in the service industry. These are the Insurance Model, where the service provider makes its best attempt to satisfy the performance, availability and responsiveness objectives that are specified in the SLA according to its normal operating procedures, the Provisioning Approach where the service provider typically signs different types of service objectives with different customers and allocates the resources within the environment differently to each customer in order to be able to support the service level objectives for each of the individual customer, and finally the Adaptive approach where service provider would dynamically modify the configuration of the system used to support the customer when monitoring mechanisms indicates a change in requirements and a danger that the SLA might get violated. Research in this paper through interaction with the industry and the academia indicates that for the cloud based service framework, all three models would be required and some customization on the model might be used at times, based on the kind of type of services desired by the broker and the consumer.

D. Constituents of an SLA for the Service Industry

In a service oriented architecture, especially on an IP based networks which the cloud paradigm is all about, the creation of an SLA would entail incorporating several system availability, system performance and security related metrics. A tentative list is provided in the GICTF [21], for ready reference. The final aim of providing a service level management framework is to enable the players to offer a business ready service oriented architecture that enables the service economy in a quantifiable and dependable way. This is true for cloud providers, consumers and brokers alike. Thus the intended SLA governing the relationship between these actors must ensure that the following metrics are met:

• The quality characteristics of service are predictable and enforced at run time.
• The SLA management is transparent and defines the exact conditions of service delivery and can be managed across the entire IT service stack as defined in the NIST model.
• The whole process is as automated as possible to ensure that the service delivery is elastic and scalable, besides being responsive.
• The process of creating an SLA must be repeatable. How this translates to a cloud broker-cloud consumer and cloud broker-cloud provider is the subject of the next section.

IV. SLA WITHIN THE CLOUD BROKER PARADIGM

As cloud computing is evolving, the provisioning and monitoring of cloud services is becoming more complex. It has been realized that the present set of services on offer are so complex that normal cloud consumers would not be able to manage and deploy them without significant assistance. In such a scenario, a cloud consumer would request cloud services from a cloud broker, instead of contacting a cloud provider directly [1]. As per NIST, a cloud broker is an entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers.

A. SLA Formulation Issues in the Cloud

Ensuring SLA formulation in the present cloud service provisioning space is a non trivial task. Compliance to multiple local laws in the location that house the data of the cloud consumer, opacity in terms of location of the resources that are provisioned and other similar non-quantifiable metrics make the drafting, measuring and monitoring difficult. The present framework of cloud provisioning is by no means stable and the interplay between players in the cloud model is presently not able to efficiently and adequately address the needs of consumers or the brokers. There is thus a growing need for adopting SLA frameworks that not only support the service models of IaaS, PaaS and SaaS, but also provide a measuring and quantification methodology for ensuring SLA adherence. This issue finds mention in the Draft NIST Roadmap for Cloud Computing, in Section 2.3, which highlights the need for an industry wide standard SLA for provisioning of services between the cloud provider and the cloud consumer. The draft is however
silent on the need to formulate the SLAs between the broker and other players in the model.

As per the NIST framework, the cloud broker would provide three distinct services: Service Intermediation, Service Aggregation, and Service Arbitrage. These have been explained in detail in the ibid document and the distinction lies in the mode of provisioning of the services and what kind of value addition the broker would provide to the cloud consumer and a business value to the cloud service provider. These require specific and binding agreements between the actors for the reference model to function as intended. Adding complexity to the cloud brokering framework are the varied deployment models that exist in reality, i.e. the public, private hybrid and community deployment models. The broker would require multiple SLAs with the associated stake holders based on the deployment model and the placement of actors in the model.

---

Figure 1. Framework for SLA management and Cascading Effect

B. Cloud Broker and Cloud Consumer SLA

Cloud service oriented SLAs, with the cloud broker as an actor, represent a negotiated service contract between the associated parties that specifies, in measurable terms, what cloud service will be provided to the consumer through the cloud broker. This necessitates that key elements required for cloud services including warranties, guarantees and related performance metrics are not left out of the SLA. If left out, they often tend to make the SLA unenforceable. The broker would need to make the consumer understand and appreciate the nuance of such elements and make sure that the agreement between the broker and the service provider also reflects the terms in an unambiguous manner. The aim is to make sure all parties understand and anticipate the course of action in provisioning of the service.

Research indicates that the usage of common terms and definitions within the SLAs are accepted to avoid misunderstandings between all three parties. The terms of reference need to be universally defined at the beginning of an SLA in a manner that it becomes unambiguous to the consumer, the provider and the broker as to what the service agreement entails.

C. Cloud Broker and Cloud Provider SLA

It is also necessary to create an environment which allows the broker to objectively compare competing services and offer them as bundles to the intended consumers. As the broker would be involved in service intermediation, aggregation, and arbitrage, it is necessary to have a comparative framework where the services provisioning and service usage are both compared in an objective manner. The authors are of the opinion that reputation based systems would be ideal to achieve such objectivity. Design of such systems would be a work in progress and evolve based on the stability of the cloud broker system.

SLAs that would define the relationship between cloud brokers and cloud providers would need to be based on the same lines as those between the cloud broker and the cloud consumers. There is a need for enumerating the same level of service provisioning guidelines which get mentioned in the broker-consumer SLA.

Fig. 1 illustrates the relationship between the actors involved in the service provisioning model and how enforceable SLAs would provide a systematic assessment of the services on offer based on the measuring, monitoring and penalty metrics. The figure also illustrates the effect of a failure of an SLA on the provisioning model. With the cloud broker as an entity in the NIST recommended cloud framework, it is imperative that metrics of service agreement agreed upon between the cloud broker and cloud service provider need to be more stringent than those between the broker and the cloud consumer. A failure in provisioning of the agreed upon services by the cloud service provider will have a ‘Cascading Effect’ on the service model. The ‘cascade’ will be aggravated in the cloud paradigm as multiple associations exist between the cloud broker and the service consumers (one-to-many and, at times, many-to-many). The aspect of service arbitrage by the cloud broker would thus need to be deliberated very minutely in the event of a failure of service. The researchers strongly believe that the affect of SLA failures will lead to a cascade effect in terms of service outage for multiple cloud service consumers. In April 2012, Amazon Inc., faced a major outage of host of its services [22]. Such outages reflect the effect of the cascade due to the failure of a bundle of services from one provider on multiple, sometimes more than a million consumers – which we feel is a cascade of service outages.

The researchers are also convinced based on interaction with the industry and the academia that there is a need for a reputation based system, based on the assumption of a stronger metric enforcement between the cloud broker and cloud service provider vis-à-vis the cloud broker-consumer, for arriving at a comparative framework for selecting the service bundle and defining the system level and availability based metrics in the SLA between the cloud service provider
and the broker. The reputation based system can be based on relevant service metrics as would be proposed for the consumer-broker SLA. Some of the metrics, which the researchers feel could be used in selecting the appropriate service bundles by the brokers could include response time in provisioning (SaaS), rate of successful delivery of promised services levels of a defined period of time (PaaS), risk preventing mechanisms in place by the provider and SLA success metrics of the provider. This also brings upon the aspect of measuring mechanisms which need to be in place while drafting the SLA. This is especially true when drafting the provider–broker SLAs as it is anticipated that aggregation of multiple, differing services is the way ahead and cloud brokers would need to have a mean to measure the service been hired. This is also illustrated in Figure 1 above. This research illustrates that the violation of service agreements between the broker and provider has a consequent affect on the agreements between the broker and the consumer and this can lead to a cascading degradation in service provisioning, if not checked in time through effective monitoring mechanisms. SLA drafting and management by incorporating effective monitoring and measuring mechanisms is thus an essential task in ensuring better cloud services provisioning. The metrics recommended in this research for basing a SLA between different actors in the cloud framework would need further study and the researchers also believe based on the work thus far that these metrics would change based on the service bundle desired by the consumer and arbitraged by the broker.

V. CONCLUSION

Creating an effective trust relationship between the cloud brokers and cloud consumers is essential to maintain the agreed level of service provisioning in the cloud. This trust is enforced using effective agreements between actors. This trust is often realized when agreements are based on clearly defined and effectively executed contract agreements, or SLAs, which are a corner stone for provision of well executed, responsive and elastic services in the cloud. The aspect of SLA management between cloud brokers and cloud providers as well as between cloud brokers and consumers is a research void at present and has been highlighted in this paper. The SLAs between the three actors have a bearing on each other. Industry reports coupled with the research done on the subject indicate that the cloud broker’s role in the framework for cloud service provisioning is increasing and thus the relationship between these individual SLA assume increased importance. It has also been realized through this research that a strong contractual SLA between the cloud broker and the cloud service provider is necessary for the cloud framework to maintain its stability. The research also highlights the affects of failure of the agreed upon services illustrated in a SLA between the broker and the provider and the consequent service outage which ensues. The researchers have termed this as ‘Cascading Effect’ in the cloud service model. The utility of a well defined and enforceable SLA based on quantifiable metrics with the broker as a central actor is thus of paramount importance. The researchers also believe that there is a need for a reputation based, comparative system for arbitraging services from different service providers. Such a system can be used by brokers for selecting the bundle of service more efficiently and the design of such a reputation based system is a work in progress.

VI. FUTURE WORK

This paper is a work in continuation as part of a doctoral thesis on cloud computing and affects on managerial aspects of an organization when working in a cloud paradigm. As a future work the authors are examining the NIST framework and exploring how measurable metrics can be defined to create a universally acceptable interoperability framework required for dissimilar clouds to talk to each other. The authors are also of the opinion that there is a need to further work on drafting comprehensive and binding SLA templates that address the lacunae existing in service provisioning between the three actors. There is also a need to further understand the cascading affect due to terms of service violation when seen from the perspective of a cloud broker. Monitoring and measuring frameworks also form an essential part of the SLA management process and are a topic for future research. Another work in future could be the affect of these SLAs on the cloud auditors and cloud carriers when viewed from the perspective of a cloud broker.

REFERENCES


