**Indian Journal of Science and Technology,** *Vol 9(48), DOI: 10.17485/ijst/2016/v9i48/109302, December 2016*
*ISSN (Print) : 0974-6846*
*ISSN (Online) : 0974-5645*

# Service Level Agreement (SLA) Enforcement: Is it Impactful?

**Rodziah Atan and Hitham Haider**

Department of Software Engineering and Information Systems, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia; rodziah@upm.edu.my, hith2001@gmail.com

## Abstract

Service Level Agreement (SLA) enforcement's impact measure is a potential research area to be explored. Assumptions that this study is making are, SLA management will become better by a firm enforcement, which monitors and encourage every customer to be responsible to launch report of bugs or mischief of services such as unsatisfactory service quality or service unavailability to a collection pool where the provider react immediately (based on agreed terms) to address the complaints, so that the total DownTime (DT) will not exceeds the agreed SLA value. This study establishes fundamental theory to measure the enforcement impact based on automatic SLA monitoring and management.

**Keywords:** Efficiency, Impact Measures, Service Level Agreement (Sla)

## 1. Introduction

SLA or abbreviated as SLA is not a new concept. It has been used since decades ago as a contract binding document between service providers and service customers. SLA contains service terms to service customers or subscribers. The term 'service' is subjective in nature. Service measurement normally use subjective terms also, such as 'satisfies', 'fulfils', 'quick', 'unhappy' or 'slow'.

Sometimes it is difficult to explain service's mischief where customers cannot find exact words to put them. There are services that can be easily measured such as courier services where for each parcel or letter, once it is delivered to respective recipient, it means that the service has been complied. Services in transportation, banking or construction normally follows a strict procedures and policies, and what concerned by customers are deliverables or outcome of the services they are getting. For transportation, service expected should be goods or people be delivered safely and in one piece. As for services such as utility services (telecommunication or health), it is a bit more complicated to explain their terms of service measurement.

SLA is employed to make service providers and service customers understand and mutually agreed to the terms of services and other necessities the providers are providing and the customers are getting. SLA is a type of contract that must be carefully read and understood before signed.

SLA implementation in digital domain has a slight different account. The SLA content mostly governs by Service Level Objective (SLO) which highly focused and tuned to specific service. An example of SLO is paid TV channel service subscription.

The SLA will contains items to be agreed such as number of channels, 24/7 subscription uptime, charges per months, early notices of maintenance work or server shutdown, etc. SLO is a set of formal expressions. These formal expressions have the well-known "if...then" structure. The antecedent "if" contains conditions and the consequent "then" contains actions. An action represents what a party has agreed to perform when the conditions are met. For example, the maximum subscription uptime should be 24 hours 7 days a week for 45 channels, and should not be below certain specific value. On receiving the values, an evaluation service evaluates the SLO predi-

**Author for correspondence*

cates. In case of SLA violations, management service is notified to process the penalties.

In most of literatures related to SLA monitoring, centralized monitoring system technique has always been employed, either in provider, consumer or trusted party sides. Mostly, the SLA management system and monitoring system are separated; therefore, it creates space from having autonomic and self-managed SLA. Examples are such as, LoM2HiS[1] monitoring framework which is attached in FoSII[2] infrastructure, utilizing the metrics value from host monitoring, which then calculated the SLA attributes based on proposed formula by[3]. All SLAs and their features in this framework consist of attributes, metrics and formulas that have to be located in central repository; which is another monitoring system where it is then used to evaluate the SLAs.

Similarly, SLA monitoring in provider side is presented, where they normally are installed in virtual machine[4]. QoSMONaaS approach in[5] put concerns to the QoS monitoring. Although this approach is trusted third party SLA monitoring side, it still requires external service for monitoring and validating its results and status. LAYSI framework, proposed is one of the few researches that concentrated in layered monitoring structure[6]. This framework distributes the monitoring process and SLA management into different actors.

## 2. SLA Establishment and Implementation

A well-written SLA ensures responsibilities and benefits of services for both providers' and customers' sides. The building blocks of an SLA consists of items such as assesses to current operations and services, and the service levels.

Investigating the current situation of SLA management shows difficulties in fully implement SLA monitoring and its related management due to several affecting factors. Current SLAs have the agreement in place and they should be followed as agreed. However, implementing agreed SLA still need to be improved. Enforcement of SLA total application is seen to be the center focus. Therefore, measuring the efficiency of SLA enforcement at certain level of SLA implementation would be the first step possible to increase SLA monitoring and management efficiency.

## 3. Service Level Definition

The efficiency measurement theory for SLA enforcement is proposed to study the impact and factors affecting SLA management efficiency. The theory adapted SMART theory establishment technique[1]. SLA normally contains information such as record of terms of the agreement, outlines of the roles and responsibilities for the customer and the service provider, contract duration, locations of service and service times. Example of SLA contents are:

1. Duties of the service provider
2. Duties of the customer
3. Responsibilities of service users
4. IT Security aspects to be observed

Exceptions to service times and others are also being included in the SLA such as weekends and public holidays as well as regular maintenance downtime. Information stating about the minimum and expected performance levels for the service as well as conditions under which the service is considered to be unavailable or limited such as, the expected and minimum service levels might be 96% and 86% on schedule. The key here is that the "expected level" is what the customer is actually paying for and the "minimum level" is what the customer would consider poor services.

Escalation procedures are outlined to define steps to be taken when service levels do not meet the expected and agreed-upon standards. Other important definition in SLA is the service metrics definitions and certain way to track them over time. These items are such as the conditions when the service is considered to be unavailable/limited, availability targets, reliability targets, time-to-restore service and maintenance downtime. Commonly used metrics include:

- MTBF - Mean Time Between Failures
- MTBSI - Mean Time Between Service Incidents
- MTRS - Mean Time to Restore Service
- TAT – Turn Around Time
- Uptime, etc.

Other than those specified above, the fees and conditions applied and the exact circumstances under which they applied, where clearer conditions are stated, and the lower likelihood for disagreement. This will result in higher customer satisfaction and more prompt payment.

Penalty caps and termination of service contract termination are also important in every SLA.

A good SLA will help organization to promise what is possible to deliver and deliver what is promised. SLAs are not a commitment to deliver the impossible. A SLA can be as informal as a performance target or as rigid as a committed time to restore a system to operation backed by penalties. In either case, the SLA serves as a basis for establishing a shared understanding of the service relationship. When properly developed, SLAs offer a win-win situation for both the service provider and the customer.

## 4. Establishing Theory for SLA Efficiency Measures

The proposed theory is to find the inter-relationship between service providers' reactions towards certain categories of SLA violations through stated set of activities with service customers' SLA violation discern and feedback.

Some outlined research questions are such as what are the factors involved in certain SLA management and enforcement? How violation is detected? What will be the earliest respond towards SLA violations from both provider and customer sides? and Who is responsible to monitor SLA? This theory holds to certain key concepts of theory establishment such as the enforcement impact to SLA management, SLA monitoring and management, detected SLA violation penalty calculation, SLA enforcement through distribution of respond and feedback responsibilities

### 4.1 Assumptions and Hypothesis

Assumptions that this study is making are: SLA management will become better by a firm enforcement, where every customers are responsible to launch report of bugs or mischief of services (unsatisfactory quality, unavailability) to a collection pool, by which whom will issue tickets (complaint tokens) to provider to fix the services' bugs or mischief. The provider side, once received the token will react immediately to the complaints so that the total DT (time of launched complaints until the bugs or mischief is fixed) is not exceeding (or below) the SLA value. If the total DT exceed (or below) the SLA value, penalty will be calculated and the provider will have to pay (or give out rebates or discounts) to the complained customers.

Model to be applied for efficiency measure and decision support will be the Simple Multi Attributes Rating Technique (SMART) model to calculate the most effective approach[7] for efficiency of services measure in the context of SLA monitoring and management. The model analysis should able to produce realistic scores and tradeoffs for SLA efficiency measures.

Hypotheses of this study are indicated as follows:

- H0 → Relating factors of SLA management and monitoring DO NOT have any INFLUENCE towards SLA efficiency measures
- H1 → Relating factors of SLA management and monitoring HIGHLY INFLUENCE SLA efficiency measures

### 4.2 Information Measurement

Information measurement is divided into two sections, Attributes for Fx = (H, 3P, E) and Attributes for Fy = (L, Q):

- Fx attributes contains three factor, which are H (Human), 3Ps (Policy, Procedure and Process) and E (Enforcement). The attribute scales are defined as such that these factors could indicate the *Actors* (Fx) in the study.
- Fy attributes contains two factors, L (Launched Complaints) and Q (Frequency of monitoring). The attributes scales are defined as such that these factors could indicate the *Events* (Fy) in this study.

The relationships of these two attributes will form the hypotheses set for *Efficiency* (EFxy) of services from the perspective the end users (either from Provider side or Client side).

### 4.3 Attribute Scales - Actors (Fx)

There are three attribute scales defined for *Actors* in this theory, as described below:

#### *4.3.1 Human (H) (e.g. Productivity)[Def. 1] Scale*

*[*-1]User group productivity is diminished sufficiently that noticeably longer time or more resources are required to provide the same level of service.

[0]No change in user group productivity is perceived.

[1] User group productivity is enhanced to the extent that group members are perceived by their clients to be providing better service, or somewhat fewer resources are required to provide service at the same level as before the network was installed.

[2]There is significant and easily perceived increase in user group productivity. Indicator of this could include a significant reduction in the staffing level required to carry out user group activities or considerable improvement in the financial performance of group.

Although subjective, the design team is sure that they can rate the productivity enhancement from each alternative network design according to these scales;

### *4.3.2 Policy/Procedure/Process (3Ps) (e.g. Security) [Def.2] Scale*

[-1]The addition of the network causes a potentially serious decrease in system control and security for the use of data or software.

[0]A noticeable but acceptable diminishing of system control and security

[1]No detectable change in system control or security is enhanced by addition of a network.

### *4.3.3 Enforcement (E) (e.g. Compound Fees) [Def. 3] Scale*

[-1]No charges fee issued to service violation environment

[0]Medium amount of charges fee issued to service violation environment for total unavailability of services

[1]Large amount of charges fee issued to service violation environment for total unavailability of services

## 4.4 Attribute Scales - Events (Fy)

There are two attribute scales defined for *Events* in this theory, as described below:

### *4.4.1 Launched Complaints (L) (e.g. Number of Complaints, Contento of Complaints) [Def. 4] Scale*

[-2] No launched complaints for a year of services.

[-1] User launched complaints rarely (>20 complaints for a year of services) and stating straight forward complaints (shows that the user understands how the services should be – easily understandable 3Ps).

[0] User launched complaints rarely (>20 complaints for a year of services) and stating complicated complaints (shows that the user does not really understand how the services should be).

[1] User launched complaints often (>40 complaints for a year of services) and stating straight forward complaints.

[2] User launched complaints often (>40 complaints for a year of services) and stating complicated complaints.

### *4.4.2 Frequency of Monitoring (Q) (e.g. Regular Service Check) [Def.5]*

[-1] No service checks and maintenance works scheduled within a year of services

[0] No maintenance works scheduled and service checks are only done based on launched user complaints.

[1] Maintenance works is scheduled with long duration gaps and service checks only done based on launched user complaints.

[2]Maintenance works is scheduled with short duration gaps and maintenance works only done based on launched user complaints.

[3]Frequent service checks and maintenance works is scheduled with short duration gaps within a year of services, launched user complaints are treated priorities.

## 4.5 Theory Inference, Interpretation and Evidence

This research utilizes theory inference to derive logical conclusion from all attributes considered, i.e. H, 3P, E, L and Q. Deriving logical conclusion from the observed measurement attributes of *Fx* and *Fy* in this context of study, where, the inferred conclusion should depict the relationships of service efficiency ($Ef_{xy}$) with attributes for each factors (*Fx* and *Fy*). Inductive inference is used in this study. Inductive inference allows the attributes sets in this study to supply strong evidence to form conclusions. Meaningful relationships are articulated, based on interviews, observations, documentations and logic evidences gathered for this study. The main conclusion of this study will be based on accumulated articulation of meaningful relationships generated. Statistical inference using mathematical equations and notions will be employed to draw conclusions in the presence of uncertainty (i.e. factors of culture, type of government services, human attributes and behavior) tackled by this study.

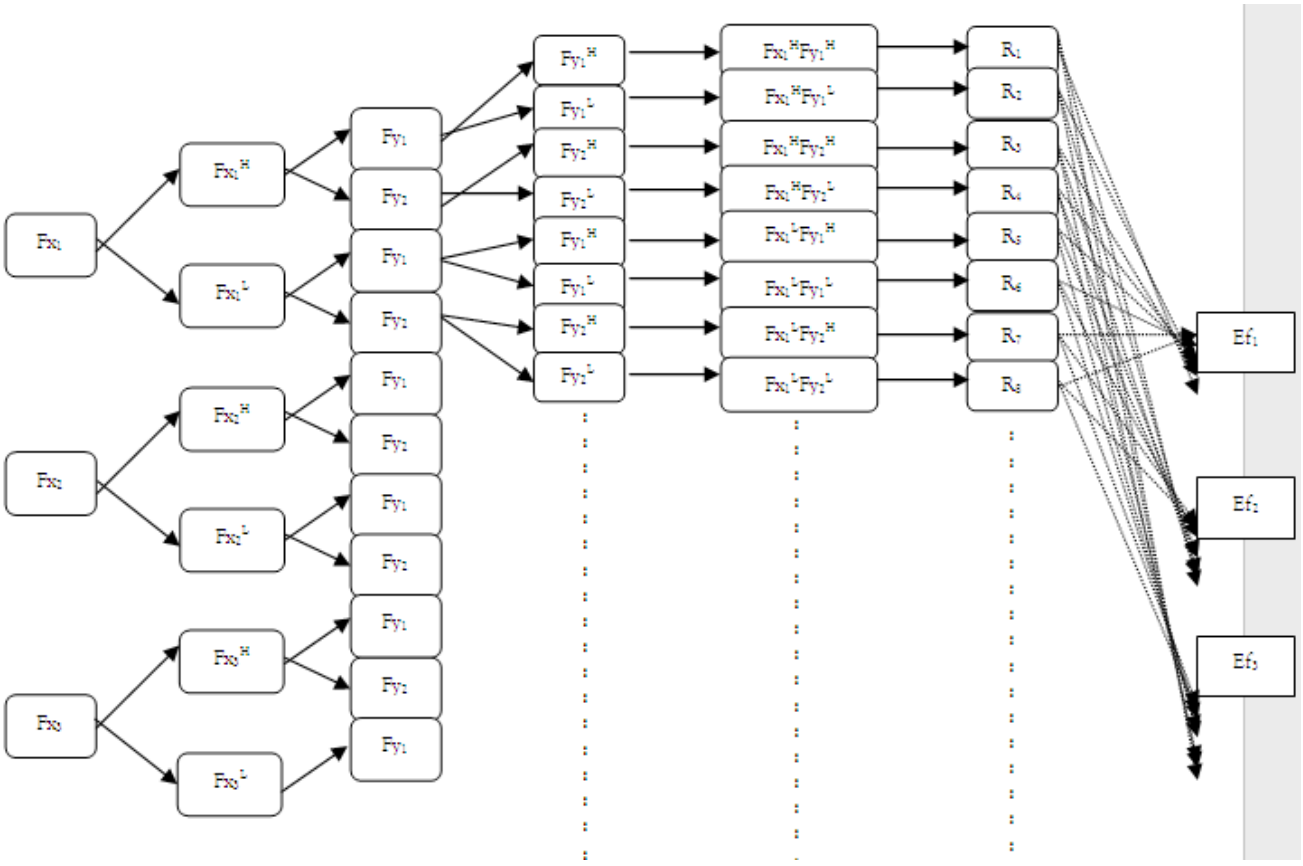


**Figure 1.** Factors relationships and efficiency implies model.

However, in this study, we do our best to apply inferential statistical analysis inference properties for a sample population for hypothesis testing and driving estimates from it. The analysis of this study establishes Relationships ($R_{1..n}$). Factors of *Fx* ($Fx_1$, $Fx_2$, $Fx_3$) and *Fy* ($Fy_1$, $Fy_2$) implies to Efficiency ($Ef_1$, $Ef_2$, $Ef_3$) of services based on specified SLA. Figure 1 shows the relationships and efficiency implies from the study. Efficiency in this study is divided into three categories:

- $Ef_1$ – low efficiency, by which, relating factors indicate low efficiency of services monitoring and management which implies to low quality of service and low collection of service violation compound fees (or high rebates issued).
- $Ef_2$ – medium efficiency, by which, relating factors indicate medium efficiency of services monitoring and management which implies to medium quality of service and medium level collection of service violation compound fees (or medium rebates issued).
- $Ef_3$ – high efficiency, by which, relating factors indicate high efficiency of services monitoring and management which implies to high quality of service and high collection of service violation compound fees (or low rebates issued).

The proposed three *Fx* factors and two *Fy* factors are most important in this study. The factors might be changeable in terms of numbers and levels to follow suit own context. For each of factor applied, possible relationships can be expected as the pushing and pulling factors affecting efficiency of SLA measurement.

Analysis Method used SMART is based on a linear additive model, which means that an overall value of a given alternative is calculated as the total sum of the performance score (value) of each criterion (attribute) multiplied with the weight of that criterion. This technique is adapted to equip decision making from service recipient side (in this context, the government agencies) based on proposed efficiency measurements model.

## 5. Conclusion

This research proposed a theory to measure the impact of SLA enforcement to both service receivers and provid-

ers sides. The theory aimed to facilitate the SLA execution enforcement by focusing on affective factors based on specific SLA management cases. Once the SLA enforcement efficiency measures theory is able to tell the affecting factors, this theory can be applied to any service monitoring and management activities throughout the agencies. This approach should able to minimize effort in identifying problematic factors in SLA management.

## 6. Acknowledgment

This work is funded by Malaysian Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS). MOHE Reference Code: FRGS/1/2015/ICT04/UPM/02/2, Project No.: 08-01-15-1274FR to increase SLA enforcement efficiency in this country.

## 7. References

1. Emeakaroha VC et al. Low level metrics to high level SLAs-LoM2HiS framework: Bridging the gap between monitored metrics and SLA parameters in cloud environments. International Conference on High Performance Computing and Simulation (HPCS). Institute of Electrical and Electronics Engineers (IEEE); 2010. p. 48–54.
2. Foundation of self-governing ICT infrastructures (FoSII) [Internet]. 2012 [cited 2012 Mar]. Available from: http://www.infosys.tuwien.ac.at/linksites/FOSII.
3. Emeakaroha VC et al. Cloud resource provisioning and SLA enforcement via LoM2HiS framework. In Concurrency and Computation Practice and Experience, Wiley Online Library; 2012 Oct 10.
4. Al-Ghuwairi AR, Cook J. Modeling and enforcement of cloud computing service level agreements. Technical Report; 2012. p. 1–9.
5. Cicotti G et al. How to monitor QoS in cloud infrastructures: the QoSMONaaS approach. Intelligent Distributed Computing VI. 2013; 446:253–62.
6. Ul Haq I, Brandic I, Schikuta E. (2010). SLA validation in layered cloud infrastructures. Economics of Grids, Clouds, Systems, and Services, Springler, Lecture Notes in Computer Science. 2010; 6296:153–64.
7. Barfod MB, Leleur S, editors. Multi-criteria decision analysis for use in transport decision making. In the 2 edition of DTU Transport Compendium Series, Technical University, Denmark; 2014. p. 5–75.