
A Novel Approach Based on Service Level Agreement (SLA) for Evaluating the Web Service Quality

Jamila Qaderi Aziz

Ph.D. of Business Administration, Faculty of Business, Ain Shams
University, Egypt
jamilaazzizq@yahoo.com

Aref Sheriff Banhawi

Ph.D. of Information Technology, Faculty of Computers and Information, Zagazig
University, Egypt
sheriff.aref@gmail.com

Abstract:

Web service technology has gained a more important role in developing distributed applications and systems on the Internet. The rapid growth of published web services makes their discovery more and more difficult. Nowadays, most web service providers sign Services Level Agreement (SLA) contracts with their clients in order to guarantee the offered functionality of their services. This paper proposes an approach to monitor the Quality of Services (QoS) in web service according to Service Level Objectives (SLO) in SLA. Monitoring procedures are introduced to check variations in the pre-agreed metric values of SLAs. Then, the deviation between the actual quality and the acceptable quality level can be identified and analyzed. Finally, the weaknesses of the web service practices can be discovered and solved.

Keywords: Web Services, Services Level Agreement, Quality of Services, Web Service Level Agreement, Evaluation.

1- Introduction

A web service is a software system designed to support interoperable machine-to-machine interaction over a network [1,2, 3]. Other systems interact with the web service using Simple Object Access Protocol (SOAP) messages, typically conveyed using HTTP with an XML serialization in conjunction with other web-related standards [3,4, 5]. The web services architecture is based upon the interactions between three primary roles: service provider, service registry, and service requestor [6]. These roles interact using publish, find, and bind operations [7]. The service provider is the business that provides access to the web service and publishes the service description in a service registry [7]. The service requestor finds the service description in a service registry and uses the information in the description to bind to a service [7, 8]. Service registry is a searchable registry of service descriptions where service providers publish their service descriptions [7].

Web services are composed of functional and non-functional attributes. The non-functional attributes are referred to as QoS. QoS is defined in [9], adapting from the definition of quality in ISO 8402, as a set of non-functional attributes of the entities used in the path from a web service repository to the consumer who relies on the ability of a web service to satisfy its stated or implied needs in an end-to-end fashion. Some examples of QoS attributes are performance, reliability, security, availability, usability, discoverability, and adaptability [10, 11].

A QoS attribute of a web service may have a metric for quantification that can be used to assess the service's performance [12]. This metric is specified in the web service contract between service partners. The measure provides a quantitative indication of the extent, amount, dimensions, capacity, or size of some attributes of a product or process [13]. Monitoring is the process of observing the quality of the service over a period and finding the degree of deviation from the expected norm

[14]. In the case of monitoring statically composed services, corrective actions for SLA violations can be incorporated only by stopping the execution and resuming it with an alternate concrete composition [14].

SLA is a contract between a network service provider and a customer that specifies what services the network service provider will provide. Many Internet Service Providers ISPs provide their customers with an SLA [8]. SLA faces many challenges such as a lack of accurate input, involvement, and commitment from the business and customers, the tools and resources required to agree, document, monitor, report, and review agreements and service levels, the process becomes a bureaucratic, administrative process rather than an active and proactive process delivering measurable benefit to the business, business, and customer measurements are too difficult to measure and improve, so are not recorded, inappropriate business and customer contacts and relationships are developed, high customer expectations and low perception, poor and inappropriate communication are achieved with the business and customers [15].

This paper proposes a proposed approach for monitoring the QoS of SLA in web services. This approach used event-based monitoring and history-based monitoring. It includes many activities within SLA and checking actual metrics with pre-defined to detect variations.

2- Related Work

Many researchers in this domain focus on the main points: monitoring of web services, SLA management including QoS management, and mapping techniques of monitored metrics to SLA parameters and attributes.

Karthikeyan. J et al, propose a framework called Quality based Dynamic Composition (QDC) framework. This framework includes ontology for improving

the web semantic discovery and also provides the best service according to the customer specification as well as monitoring the execution of web service and evaluating the QoS parameters associated with web service [8]. Katawut Kaewbanjong et al, review attributes, categorization, and metrics of QoS of web services. In addition, create a new classification scheme where each category under this scheme can point to the information needed for the successful application of QoS attributes [13]. Shanmuga Priya R et al, propose an approach to monitor the QoS properties of compositions that are dynamically composed. Aspect Oriented Programming (AOP) based approach is adapted for monitoring and triggering alternate service composition, in case of SLA violations. The proposed approach is demonstrated for an application of the foreign exchange market that involves composing services at runtime [14]. G. Dobson et al, present a unified QoS ontology applicable to the main scenarios identified such as QoS-based web services selection, QoS monitoring, and QoS adaptation [16].

S. Ran proposes a QoS model and a Universal Description, Discovery, and Integration (UDDI) extension for associated QoS to a specific web service. The author does not specify how QOS Model values are assessed and monitored. It is assumed that they are specified by the service provider in UDDI [17]. M. Comuzzi, proposes an architecture for monitoring SLAs considering two requirements introduced by SLA establishment: the availability of historical data for evaluating SLA offers and the assessment of the capability to monitor the terms in an SLA offer [18].

H. G. Song et al, propose a simulation-based web service performance analysis tool that calls the web service once under low load conditions and then transforms these testing results into a simulation model [19]. D. Gunter et al, present a NetLogger that is a distributed monitoring system, which monitors and collects information from networks. Applications can invoke NetLogger's API to survey the overload before

and after some request or operation. However, it monitors only network resources [20].

T. Suzumura et al, work on performance optimizations of web services. Their approach is to minimize XML processing time by using differential deserialization [21]. B. Koller et al, discuss autonomous QoS management using a proxy-like approach. SLAs can be exploited to define certain QoS parameters that a service has to maintain during its interaction with a specific customer. However, their approach is limited to web services and does not consider the requirements of cloud computing infrastructures like scalability [22]. M. Tian et al present an approach for integrating QoS with web services. The authors implemented a tool suite for associating, querying, and monitoring the QoS of a web service [23]. I. Brandic et al present an approach for adaptive generation of SLA templates. SLA users can define mappings from their local SLA templates to remote templates to facilitate communication with numerous cloud service providers. However, they do not investigate the mapping of monitored metrics to agreed SLAs [24].

I. Brandic et al deal with QoS attributes for web services. They identified important QoS attributes and their composition from resource metrics. They presented some mapping techniques for composing QoS attributes from resource metrics to form SLA parameters for a specific domain. However, they did not deal with monitoring of resource metrics [25]. F. Rosenberg et al and L. Zeng et al, propose a QoS model and a middleware approach for dynamic QoS-driven service composition. They investigate a global planning approach to determine optimal service execution plans for composite services based on QoS criteria [26,27]. L. Zeng et al, introduce a model-driven approach for integrating performance prediction into service composition processes carried out using BPEL. They composed service SLA parameters from resource metrics using some mapping techniques. However, they did neither consider resource metrics – nor SLA monitoring [28].

3- Service Level Agreement

SLA sets the expectations between the consumer and provider [29]. It helps to define the relationship between the two parties. It manages how the service provider sets and maintains commitments to the service consumer [6]. A properly specified SLA describes each service offered and addresses:

- How the delivery of the service at the specified level of quality will become realized.
- Which metrics will be collected?
- Who will collect the metrics and how?
- Actions to be taken when the service is not delivered at the specified level of quality and who is responsible for doing them.
- Penalties for failure to deliver the service at the specified level of quality.
- How and whether the SLA will evolve as technology changes (e.g., multi-core processors improve the provider's ability to reduce end-to-end latency) [30]

SLAs can be either static or dynamic. A static SLA is an SLA that generally remains unchanged for multiple service time intervals. A dynamic SLA is an SLA that generally changes from service period to service period, to accommodate changes in the provision of service. Any SLA may contain the following parts: purpose, parties, validity period, scope, restrictions, service-level objectives (availability, performance, and reliability), optional services, and administration authority [31]. In the following sub-sections, the researchers explain SLA importance, the SLA life cycle, and how to formulate SLAs in web services.

3-1 SLA Importance:

SLA is important [32] because it sets boundaries for the following aspects of service provisioning.

- Customer commitments: Focused on customer requirements and ensured that the internal processes followed the right direction.
- Key performance indicators for customer service: Improved customer satisfaction is a clear objective.
- Key performance indicators for the internal organizations: Internal objectives become clearer and easier to measure.
- The price of non-conformance: If the SLA has penalties non-performance can be costly. However, by having penalties defined, the customer understands that the provider truly believes in its ability to achieve the set performance levels.

3-2 SLA Life cycle

SLAs have a certain life cycle that consists of six phases [33]:

1. Service and SLA Template Development: This phase includes the identification of (service consumer needs, appropriate service characteristics and parameters) that can be offered given the service execution environment and the preparation of standard SLA templates.
2. Negotiation: This phase includes the negotiation of the specific values for the defined service parameters, the costs for the service consumer, the costs for the service provider when the SLA is violated, and the definition and periodicity of reports to be provided to the service consumer.
3. Preparation: The service is prepared for consumption by the service consumer. This phase may require the reconfiguration of the resources that support service execution to meet SLA parameters.
4. Execution: This phase is the actual operation of the service. It includes service execution and monitoring, real-time reporting, service quality validation, and real-time SLA violation processing.
5. Assessment: This phase has two parts:

- Assessment of the SLA and the QoS that is provided to an individual consumer. QoS, consumer satisfaction, potential improvements, and changing requirements are reviewed periodically for each SLA.
 - Assessment of the overall service. This assessment can be tied to an internal business review. Elements to be covered in this review are the QoS provided to all consumers, the need for the realignment of service goals and operations, the identification of service support problems, and the identification of the need for different service levels.
6. Termination and Decommission: This phase deals with termination of the service for reasons such as contract expiration or violation of contract. Also, this phase deals with decommissioning of discontinued services.

3-3 Expressing SLAs in the Web Service Level Agreement Language

WSLA [34] was developed by IBM and is used to define SLA documents. A WSLA is an agreement between a service provider and a customer and as such defines the obligations of the parties involved. It is the obligation of a service provider to perform a service according to agreed-upon guarantees for the service parameters on the technical level. The design goals of WSLA are a formal and flexible XML-based language for SLA definitions between different organizations. WSLA is structured in three sections: the Parties section, the Service Description, and the Obligations section. The Parties section [35] identifies the contractual parties and contains the technical properties of a party, i.e., their address and interface definitions. The service description section [35] specifies the characteristics of the service and its observable parameters as follows: For every service operation, one or more is binding. In addition, one or more SLA Parameters of the service may be specified. Obligations Section [35] defines the SLOs, guarantees, and constraints that may be imposed on the SLA parameters. This allows the parties to unambiguously define the

respective guarantees they give each other. The WSLA language provides two types of obligation are Service Level Objectives represent promises with respect to the state of SLA parameters and Action Guarantees are promises of a signatory party to perform an action. This may include notifications of service-level objective violations or invocation of management operations.

4- SLA Qualities

There are two categories of qualities that can be specified in SLAs [31, 33]: measurable and unmeasurable.

4-1 Measurable Qualities

Measurable qualities can be measured automatically using metrics; for example, the percentage of time a system is available. Measurable qualities include accuracy, availability, capacity, cost, latency, provisioning-related time, reliable messaging, and scalability.

- *Accuracy*: is concerned with the error rate of the service. It is possible to specify the average number of errors over a given period.
- *Availability*: is concerned with the mean time to failure for services. It is possible to specify:
 - The system's response when a failure occurs.
 - The time it takes to recognize a malfunction.
 - How long does it take to recover from a failure?
 - Whether error handling is used to mask failures
 - The downtime necessary to implement upgrades (may be zero)

- The percentage of time the system is available outside of planned maintenance time.
- *Capacity*: the number of concurrent requests that can be handled by the service in a given time period.
- *Cost*: is concerned with the cost of each service request. It is possible to specify:
 - The cost per request.
 - The cost is based on the size of the data.
 - Cost differences related to peak usage times.
- *Latency*: is concerned with the maximum amount of time between the arrival of a request and the completion of that request.
- *Provisioning-related time*: (e.g., the time it takes for a new client's account to become operational).
- *Reliable messaging*: is concerned with the guarantee of message delivery. It is possible to specify:
 - How message delivery is guaranteed (e.g., exactly once, at most once)
 - Whether the service supports delivering messages in the proper order
- *Scalability*: is concerned with the ability of the service to increase the number of successful operations completed over a given time period. It is possible to specify the maximum number of such operations.

4-2 Unmeasurable Qualities

Unmeasurable qualities are those that cannot be measured automatically from a given viewpoint; for example, determining the cost of changing a service is difficult to automate. Measurable qualities include interoperability, modifiability, and security.

- *Interoperability*: is concerned with the ability of a collection of communicating entities to share specific information and operate on it according to an agreed-upon operational semantics.
- *Modifiability*: is concerned with how often a service is likely to change. It is possible to specify how often the service's.
 - Interface changes
 - Implementation changes
- *Security*: is concerned with the system's ability to resist unauthorized usage, while providing legitimate users with access to the service. It is also characterized as a system providing non-repudiation, confidentiality, integrity, assurance, and auditing. It is possible to specify the methods for
 - Authenticating services or users
 - Authorizing services or users
 - Encrypting the data

5- The Proposed Model

An official agreement between the service provider and the user is required to guarantee the defined level of the web service performance based on service quality factors. Such a service level agreement may be very comprehensive and at the same time very specific. The service level agreement may include the procedures to be followed by the provider and user in the case when either party fails to follow the agreement.

The researchers propose an approach for monitoring the QoS of web service using SLA between consumer and provider. This approach used Event-based monitoring and history-based monitoring.

Event-based monitoring is listening to the events in parallel to the execution of a business process to verify the nonfunctional qualities of a service. History-based monitoring is an extension to event-based monitoring by collecting events in a history event repository. It is possible to recognize QoS requirements that deal with a history of process executions. An example of such a requirement could be “eighty percent of the times a process goes into execution it must complete within one minute”.

The proposed approach includes the following activities:

1. Open SLA document.
2. Parse SLA document.
3. Extract Service Level Objective (SLO).
4. Extract metrics values from the monitoring mechanism.
5. Compute additional metrics.
6. Detect Variant finder between actual & predefined value.
7. Extract QoS variation report.

5-1 Open SLA document

In this approach, evaluation procedures run on consumer's service. The values of the SLA parameters are input for the evaluation procedure, which can run on:

- Either the service consumer or service provider
- Both the service consumer and service provider

5-2 Parse SLA document

As described in the section, a WSLA document divides an SLA into three sections: parties, service description, and obligations as described in section 2. The WSLA language is based on XML; it is defined as an XML schema then WSLA document is parsed to save it.

5-3 Extract Service Level Objective (SLO)

The obligations define the SLO, that is, guarantees and constraints that may be imposed on the SLA parameters. SLA parameters define the metrics values that were agreed between the service provider and the client. The following elements of the WSLA document [30] are collected and saved to the database:

- The web service name
- The obliged party is the name of a party that is in charge of delivering what is promised in this guarantee.
- The start of the validity period
- The end of the validity period
- The predicate that applies for the specific value
- The name of the SLA parameter
- The agreed value of the SLA parameter and
- The Evaluation Event defines when the expression of the service level objective should be evaluated.

5-4 Extract metrics values from the monitoring mechanism

Monitor web service to collect all the metrics that are available and composed metrics, such as minimum response time, average response, Total number of authentication failures, total number of authentication successes, total number of requests that caused SOAP faults, and Total number of successful invocations of the method. All these metrics (monitoring metrics and composed metrics) are saved in the database to make History-based monitoring.

5-5 Compute additional metrics

There are some other metrics that can be calculated using the metrics collected from monitoring web services such as availability. Availability [37] is the probability that a service is up and running. It can be calculated in the following way:

$$\text{Availability} = (\text{MTBF} / (\text{MTBF} + \text{NTTR})) * 100$$

Mean Time Between Failures (MTBF) [37] is the average time that a web Service can perform its agreed Function without interruption. Mean Time to Repair (NTTR) [37]: The average time taken to repair a web after a Failure.

5-6 Detect Variant finder between actual & predefined value

This step is responsible for comparing between SLO (predefined value) and monitoring the SLA parameters at runtime (actual value). Once the variant is detected, the variant does not fulfill the system quality requirement; the validator checks if this variant is accepted or not. If the variant is not accepted, then the validator adds this item to the report. This step is made for each SLO.

5-7 Extract the QoS variation report

After the SLO metrics are finished, the validator extracts a report about the QoS variant and sends it to corrective management to carry out actions. The Management Service will retrieve the appropriate actions to correct the problem, as specified in the SLA to improve the quality of service.

6- Conclusion

This paper aimed to propose an approach for monitoring QoS using SLA in web Services. Therefore, the researchers studied many attempts at monitoring of web services, SLA management including QoS management, and metrics of monitoring QoS. Then, the researcher proposed an approach that consists of seven steps: open the SLA document, parse the SLA document, extract service level objective, extract values of metrics from the monitoring mechanism, compute additional metrics, detect variants between actual & pre-defined value, and finally extract QoS variation report. The QoS variation report can help in achieving higher quality for web services. The researchers conclude that special attention must be given to SLA in dealing with the quality of web services.

In the future, the researchers aim to apply the proposed approach using real data of web services and compare the results with other results produced from other evaluation techniques of web services quality. In addition, the researchers aim to enhance the proposed approach by examining more formulas for all possible metrics, ranking, and classification of the metrics to extract the most important set, and/or utilizing the good practices of other evaluation techniques.

References

- [1] Seong-Hoon Lee, “A Study on Web Service Analysis and Bio-information based Web Service Security Mechanism”, International Journal of Security and Its Applications, Vol.8, No.2, 2014, pp.77-86.
- [2] Qusay H. Mahmoud, "Service-Oriented Architecture (SOA) and Web Services: The Road to Enterprise Application Integration (EAI)", April 2005, [online] Available: <http://www.oracle.com/technetwork/articles/javase/soa-142870.html>, [accessed 30/12/2022].
- [3] W3C, Web Services Architecture, <http://www.w3.org>, (2003).
- [4] F. Curbera, “Unraveling the Web Services Web: An Introduction to SOAP, WSDL, and UDDI”, IEEE Internet Computing, Vol. 6, Issue 2, p.p. 86 – 93, August 2002.
- [5] G. Alonso, F. Casati, H. Kuno and V. Machiraju, “Web Services Concepts, Architectures and Applications Series: Data-centric Systems and Applications”, Addison-Wesley, (2002).
- [6] Demian Antony D’Mello and Ananthanarayana V S, “Dynamic Web Service Composition Based on Operation Flow Semantics”, International Journal of Computer Applications, Vol. 1, No. 26, p.p. 1-9, February 2010.
- [7] H. Kreger, “Web Services Conceptual Architecture (WSCA1.0)”, Published May 2001, [online] Available: www.ibm.com/software/solutions/webservices/pdf/wsca.pdf [accessed: 2/1/2022].
- [8] Karthikeyan. J and Suresh Kumar. M, “Monitoring QoS parameters of composed web services”, International Conference On Information Communication And Embedded Systems (ICICES), 2014.
- [9] K. Kritikos and D. Plexousakis, “Requirements for QoS-based web service description and discovery,” IEEE Trans. on Service Computing, vol. 2, no. 4, pp. 320-337, October-December 2009.
- [10] Z. Balfagih and M. F. Hassan, “Quality model for web services from multi-stakeholders perspective”, in Proc. Information Management and Engineering Conference, Kuala Lumpur, 2009, pp. 287-291.

-
- [11] M. Marzolla and R. Mirandola, "QoS analysis for web service applications: A survey of performance-oriented approaches from an architectural view point", Technical Report, February 2010.
- [12] Mohamad Ibrahim Ladan, " Web Services Metrics: A Survey and A Classification", International Conference on Network and Electronics Engineering, IPCSIT vol.11, 2011.
- [13] Katawut Kaewbanjong and Sarun Intakosum, "QoS Attributes of Web Services: A Systematic Review and Classification ", Journal of Advanced Management Science, Vol. 3, No. 3, September 2014.
- [14] Shanmuga Priya R1, Kanchana R2, "AOP Based QoS Monitoring of Dynamic Web Service Compositions", IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2014.
- [15] Pierre Bernard, "Foundations of ITIL", Van Haren Publishing, Zaltbommel, 2012, PP. 171.
- [16] G. Dobson and A. Sanchez-Macian, "Towards Unified QoS/SLA Ontologies". Proceedings of the IEEE Services Computing Workshops (SCW 2006), 2006.
- [17] S. Ran, "A model for web services discovery with QoS", SIGecom Exchanges, 4(1):1–10, 2003.
- [18] M. Comuzzi, C. Kotsokalis, G. Spanoudkis, R. Yahyapour, "Establishing and Monitoring SLAs in Complex Service Based Systems", IEEE International Conference on Web Services 2009.
- [19] H. G. Song and K. Lee, "sPAC (Web Services Performance Analysis Center): Performance Analysis and Estimation Tool of Web Services", In Proceedings of the 3rd International Conference on Business Process Management (BPM'05), 2005, PP. 109–119.
- [20] D. Gunter, B. Tierney, B. Crowley and M. Holding, J. Lee, "Netlogger: a toolkit for distributed system performance analysis", 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2000, pp. 267-273.
- [21] T. Suzumura, T. Takase, and M. Tatsubori, "Optimizing Web services performance by differential deserialization", In Proceedings of the IEEE International Conference on Web Services(ICWS'05), 2005,pages 185–192.
-

-
- [22] B. Koller and L. Schubert, "Towards autonomous SLA management using a proxy-like approach", Multiagent Grid System, Vol.3, 2007.
- [23] M. Tian, A. Gramm, H. Ritter and J. Schiller, "Efficient Selection and Monitoring of QoS-aware Web services with the WS-QoS Framework", In Proceedings of the International Conference on Web Intelligence (WI'04), Beijing, China, 2004.
- [24] I. Brandic, D. Music, P. Leitner and S. Dustdar, "VieSLAF Framework: Enabling Adaptive and Versatile SLA Management", 6th International Workshop on Grid Economics and Business Models 2009.
- [25] F. Rosenberg, C. Platzer, and S. Dustdar, "Bootstrapping performance and dependability attributes of web service", IEEE International Conference on Web Services, 2006, pp 205-212.
- [26] L. Zeng, B. Benatallah, M. Dumas, J. Kalagnanam and Q. Z. Sheng, "Quality driven web services composition", In Proceedings of the 12th International Conference on World WideWeb (WWW'03), pages 411–421, New York, NY, USA, 2003. ACM Press.
- [27] L. Zeng, B. Benatallah, A. H. Ngu, M. Dumas and J. Kalagnanam, and H. Chang, "Qos-aware middleware for web services composition", IEEE Transactions on Software Engineering, 30(5), p.p. 311–327, May 2004.
- [28] A. D'Ambrogio, P. Bocciarelli, "A model-driven approach to describe and predict the performance of composite services", 6th international workshop on Software and performance, 2007, pp. 78-89.
- [29] Service level agreement. (2015, january) wikipedia. [Online] http://en.wikipedia.org/wiki/Service-level_agreement
- [30] L. Baresi, C. Ghezzi and S. Guinea. "Smart Monitors for Composed Services". In Procs of ICSOC'04, New York, USA, pages 308 – 315, November 2004.
- [31] Chrysostomos Zeginis, "Monitoring the QoS of Web Services using SLAs - Computing metrics for composed services", Master's Thesis, Heraklion, PP. 13-14, March 2009.
- [32] E. Wustenhoff, "Service Level Agreement in the Data Center", Sun BluePrints, April 2002.
-

-
- [33] P. Bianco Philip, G. Lewis and P. Merson, "Service Level Agreements in Service-Oriented Architecture Environments", Technical Note of Software Engineering Institute, September 2008.
- [34] A. Keller and H. Ludwig, "The WSLA Framework: Specifying and Monitoring of Service Level Agreements for Web Services", IBM research report RC22456, May 2002.
- [35] R. Kassab and Aad van Moorsel, "Mapping WSLA on Reward Constructs in Mobius", In Procs of UKPEW 2008, London, England, July 2008.
- [36] <http://www.ibm.com/developerworks/library/ws-slafram/> [accessed: 15/12/2022].
- [37] <http://www.knowledgetransfer.net/dictionary/ITIL/en/Availability.htm>
[accessed: 15/12/2022].