Literature Review

Existing literature on the quality of service (QoS) for web services revolve around a number of themes. These topical themes constituted the basis for the organisation of this review. First theme is the concept or idea of quality of service that covers what is quality of service and its development in the context of web services and the application or process of using quality of service in web services. Second theme covers the importance of quality of service in web services by identifying reasons or justifications for the engagement in quality of service. Third theme covers the issues faced by firms engaging in quality of service and the problems experienced by end-users of web services integrated with quality of service. Fourth theme is the measures and the processes of assessing web service quality through quality of service. Literature relevant to the study provides a framework for the study and identifies areas requiring further investigation.

Quality of Service (QoS) for Web Services

O'Sullivan, Edmond and Hofstede (2002) explained that web services are like other types of services having inadequate standards for use in providing an accurate description of the various facets of the service. Service standards are important because these are the factors used by service consumers as bases for actual purchasing decisions. Without sufficient service standards, consumers interested in engaging web services find it difficult to derive the information they need in making the decision because there is no clear means of comparing alternative service providers, alternative networking systems, and benefits to users. In this service environment, web service quality and web customer satisfaction are difficult to measure. Web household and business consumers target web

service quality and accessibility of information on web services comprises a quality element.

One way of deriving and providing information on web services to support perceptions of web service quality is quality of service (QoS). This process facilitates the achievement of different types of information on web services. First is discovery of information on web service quality through in a faster and reliable manner than available sources of information. Second is substitution of accurate descriptions of web service quality to provide a basis for the comparison of web service quality. Third is composition or the development of new or improved web services to facilitate optimised negotiations. Fourth is management of web services by clarifying service descriptions, identifying quality standards, assessing service quality, and controlling service delivery. (O'Sullivan, Edmond, & Hofstede, 2002)

As a concept, Menasce (2002) has described quality of service as combining a number of service-related properties. First is availability since quality of service allows users to determine the period or time of operation of the particular web service by collating information on web service operations. Second is security because quality of service is able to detect the existence of authentication mechanisms and the types of mechanisms applied in web services, data-integrity message exchanges, and resilience of web services from attacks. Third is response time since quality of service detects the response time of web services and web service-related requests. Quality of service measures response time considering arrival rates together with the number of simultaneous requests so that the response time measurement provides the mean and percentile ranking of the time it took to respond to requests. Fourth is throughput that

refers to the rate that web services process requests so that quality of service can measure maximum throughput as well as variances in the intensity of request loads. These features of quality of service is collated and organised by programs that operate by sending requests for particular types of service requests to the different web service providers. Information such as the time of processing and receipt of response together with other information are analysed to provide web service quality assessments. Evaluations of web service quality are determined by looking at both the perspectives of web service providers and web service users.

As a process, Menasce (2002) also described quality of service are the technologies used in the management of network traffic on the web using cost efficient means intended to improve the web experiences of business and home online experiences. QoS functions by enabling users to measure internet bandwidth, detect changing conditions in online networks such as the availability and traffic of networks, and direct or prioritize network traffic.

As a quality measure, Lee et al. (2003) described quality of service as assessing the web services using measures such as "performance, reliability, scalability, capacity, robustness, exception, handling, accuracy, integrity, accessibility, availability, interoperability, security, and network-related QoS requirements". The wide-ranging and comprehensive measures used by quality of service in assessing the quality of service add to the reliability of data. This means that web service providers and web service users benefit from applying quality of service or deriving web service quality data from QoS. As such, Mani and Nagarajan (2002) explained that web service users often engage in negotiations with web service providers over quality of service in web services.

As a quality-based business solution, Sumra and Arulazi (2003) provided that quality of service pertain to the functional and non-functional quality aspect of web services encompassing aspects such as accessibility, availability, integrity, interoperability, performance, reliability, and security. By looking into these different aspects, quality of service enables the organisation and management of the dynamism and unpredictability of web services. In addition, quality of service has become an area of competitiveness for web service providers by constituting an added value for web services. Web service providers need to gain an understanding of quality of service in order to adopt this as a point of competitive advantage.

As a software application, Offutt (2002) described QoS as the application of software to support web service delivery so that data on the quality measures are collected, organized and analyzed according to the quality standards applied by the QoS application user.

Existing literature has extensively discussed the different aspects of quality of service to support its potential benefits not only to the decision-making on web service providers but also on the enhancement of the management of web service provisions by web service providers. Since quality of service is multifaceted, it becomes important for interested parties to derive an in-depth understanding of quality of service in order to derive the potential benefits.

Importance of Quality of Service for Web Services

Literature relevant to the study provides a number of benefits to a number of parties. Direct benefits accrue to web service providers by deriving information on the quality of their web services based on wide-ranging criteria. Indirect benefits are

experienced by web service users comprised of businesses and households in terms of improvements in the quality of web service brought about by enhanced web services based on the problem areas determined through quality of service. These direct and indirect benefits are comprised of a number of specific advantages.

Bouch, Kurchinsky and Bhatti (2000) explained that one importance of QoS, especially to web service providers, is that this captures the criteria used by web service users or consumers in assessing the quality of web services. This means that deriving web service data on quality through QoS would allow web service providers to have an idea of the web service experiences of web service users and the extent of quality of web services. Information can become a source of strategies for service improvement. As an example, use of QoS has been able to determine that web users have the tendency not to look only at the speed of service but also at the predictability and reliability of fast web services as well as the appeal of web services and functions. Apart from using assessment criteria based on the perspective of web users, some of the criterion also directly link to the quality judgments of web users. This is exemplified by the perception of web users that web pages that allow for the fast retrieval of information are deemed more interesting when compared to websites that allow users to retrieve data at a slower rate that is also a quality of service criteria. As such, the extent of satisfaction of web users constitutes a criterion for quality service assessment using QoS.

Sumra and Arulazi (2003) added that another benefit of QoS to web service providers is as a solution to integration problems. Quality of service facilitates the easy implementation and adoption of web services based on pre-determined and explicit standards, which allows for the determination of lessons or best practices that become the

bases of web service design and service delivery improvements. As mentioned earlier, web services are complex or intricate so that the establishment of standards is difficult to accomplish. However, with an encompassing QoS, the assessment measures become the standards for web services and the service delivery targets of web service providers.

Mani and Nagarajan (2002) mentioned that still another benefit of QoS, especially to web service providers, is business success because quality of service becomes a differentiating or distinguishing factor for the many web service providers. QoS provides information on the functionality or utility of web services, which also comprise the factors determining the popularity of web service providers to web users. In addition, since quality of service comprises a success factor for web service providers, QoS becomes an important means of determining, controlling, and improving web service quality to maintain and expand the number of web users of service providers. Patronage of services then ensures revenue generation and profitability of web service providers.

Wang et al. (2004) identified another benefit of QoS to web service providers as a competitive advantage by comprising a value-added activity intended to provide fast and reliable web services to businesses and households. Porter (1998) explained that the ways of achieving competitive advantage is through cost leadership and differentiation. Wang et al. (2004) provided that web service providers could achieve competitive advantage through either of these ways by engaging in quality of service. QoS enables web service providers to save on cost from anticipating and monitoring delays in order to ease traffic build-up through web service quality monitoring data as well as achieve differentiation either by engaging in quality of service when competitors have not or by developing

innovative QoS systems to enhance data collection on the web service experiences of web users.

Wolter and Van Moorsel (2001) discussed a benefit of QoS to web service providers and web users alike, as relating to the web service provider's bottom line or the core goals of business operations. The popularity of e-services means that web servers have to ensure speed in order to ensure the satisfaction of business and household web users. Delays could mean large losses not only for the web service providers with the shifts in service providers but also for web users with business firms losing money from dissatisfied customers and households not being able to obtain e-service requests on time and according to their needs. As such, QoS covers the developing "quality of service to business" or Q2B relationships involving web service providers and web users by providing a common ground for measures to determine e-service satisfaction.

Ran (2003) discussed one benefit of QoS, particular to web service users, as providing an assurance of the application or integration of quality standards into the web services. This is an important consideration for web service users because this answers the question on the extent of quality of web services they expect to experience and on whether to engage a particular web service provider given alternatives. In addition, the use of certain types of QoS enables web users to assess whether certain web services being considered would be able to meet and match their quality requirements and whether the QoS of web service providers can be relied on.

It appears that more benefits redound to web service providers more than to users. This is because of the direct nature of the benefits accruing to web service provider and the indirect benefits to web users. The range of benefit to web service providers indicates

the need for providers to consider engaging in QoS to experience these potential advantages. Although, there is only one benefit that can be considered as directly experienced by web users, the benefits to web service provides are indirectly experienced by web users through the improvements in web services in terms of speed in response time, reliability or consistency of web service, and other aspects important to web users.

Issues in Quality of Service for Web Services

Menasce (2002) explains that in the application of QoS for web services, the emerging issues can be divided classified under the perspective of the web service provider and the service user.

In the case of web service providers, many aspects need to be considered in providing web services. One quality of service policy is 'best-effort', which means that the web service provider does not guarantee that all service requests will be processed for response so that in case of overload, pending requests will just be dropped. This QoS policy also means that the web service provider does not also make any guarantees of availability, response time, and throughput. This QoS policy works in certain arrangements for web service provision but there are also online situations that require more than just a promise that the web service provider would apply its best effort in processing requests. In instances when web service constitutes a core aspect of Internet applications covering a number of web services end users. (Menasce 2002)

In this situation, web service providers have the option to consider the policy option of engaging in long-term working relations with the users of all their services. Agreements that are more long-term could result to the establishment of service level agreements or SLAs, which are legal agreements that provides the limits of quality of

service measures so that it is clear to the provider and user the extent of the guarantees of the provider and the quality of service expected by the web user. The SLA could include specific details such as the maximum response time, the load of requests before excessive requests are dropped, and the period of service availability. (Ludwig 2003)

However, engaging in long-term agreements also involve management issues because of the unpredictability of workload shifts necessary in deriving computations as bases of compliance of the web service provider with the SLAs. In case of increased workloads or peak periods, it would be difficult for web service providers to meet all of the extents of its SLAs especially when serving multiple web users and processing many service requests. As a solution, web service providers have to apply prioritization as a control mechanism so that they may have to drop request considered as low-priority to make way for the processing of high priority requests. (Menasce 2002) Again, in prioritizing, this gives rise to the issue of the standards used to justify priority decisions (Bielski 2004).

Although, engagement in SLAs constitutes a better quality of service policy, this still involves a number of problems that affects the quality of web service delivery. As such, there is need for further enhancements of QoS measures or programs in order to assist in developing ways of enhancing web service delivery.

With regard to web service users, a number of QoS issues also emerge. One encompassing issue is the reliance upon the quality of service policy of the web service provider. A more specific issue is the management of throughout or the rate of processing service requests that affect the load of service requests. In the case of a business firm using web services to provide online services to end-users, a single website could involve

a number of related but individual services, this means that end users could send service requests for the different services and view all the services several times. This system involves a number of service requests from a single end user alone. If there are many people using the website, then the throughput increases so that in case of overloading some requests may be dropped. This is not beneficial for the business since end users could seek other websites. A solution that emerges is the bundling of these services into a single transaction so that a number of actions can be carried out as one unit. There are four types of transaction options. First is atomicity that pertains to the option of executing all actions or not at all so that users had to engage all actions to make service requests. Second is consistency so that the updates arising from a specific transaction preserve the consistency constraints. Third is isolation where transactions are not reflected until these are competed. Fourth is durability where updates on transactions that have been commenced are not lost. (Menasce 2002) There are limitations in these transactions such as the possible inconvenience for end users but when weighed with faster transaction processing could still work for web users.

Conti, Das and Shirazi (2002) mentioned four issues in the application of QoS to web services, which are: 1) the latency of document retrieval or the period between the initiation of the service request and response; 2) availability of data such as through replication; 3) utilization of network resources; and 4) congestion in networks. These have something to do with the management of service requests. A number of techniques have emerged to support the provision and experience of faster web services. One solution is caching document or locally storing data to reduce response time when re-accessed by the user. Second is prefetching that involves the fetching of other files while

the user is idle and viewing a web page. Third is pushing, which involves the saving of a bandwidth usage through the sending of data only once to many users instead of sending data to individual users singly that involves greater traffic. This carries the greatest challenge in implementation since users require different periods of retrieving data. Fourth is replication that enhances availability by balancing service access for various web servers.

These problems point to the need of enhancing quality of service applications in order to provide better web services not only to ensure the competitiveness in the industry as well as long-term relations with web users. Improving quality of service involves the determination of accurate quality measures.

Measures of Quality for Web Services

Offutt (2002) describes quality of service as a software application so that there are a number of quality measures applicable to web software, which are: 1) reliability; 2) usability; 3) security; 4) availability; 5) scalability; 6) maintainability; and 7) time to market. However, existing QoS software do not have the capability of addressing all of these measures at the same time so that combination of software are used. This gives rise to the issue of compatibility. Nevertheless, with the current rate of Internet-based developments, QoS software applications are expected to improve in the future in order to provide data on quality to facilitate improved web service delivery and monitoring.

Kim and Lee (2005) propounded a three-way quality model that involves the interrelations and interactions among the three variables quality factor, quality activity, and quality associate. Quality factor refers to the recognition of web service quality in order to usher the necessity for managing web service quality. Quality associate pertains

to the role of people and organizations as well as the tasks linked to web service delivery. Quality activity covers the different action models applied by the quality associates in managing the quality of web services. This model implies that the measures of quality of service encompasses more than the just the software application but also factors such as the people involved and the activities conducted in providing quality web services (Li 2005).

Kim and Lee (2005) further added that the measures in applying quality of service could be classified under the three variables. In the case of quality factor, the measure involved is the extent that the perception of web services as remotely used is translated into the delivery of web services to remote sites. This means that changing perspective of web services could already add to ensuring quality of web services. With regard to the quality associates, the different persons involved should collaborate so that the parties related to web service provision contributing their technical knowledge and skills in delivering quality service while web users provide valuable feedback in improving web services (Bouch, Kurchinsky & Bhatti 2000). Quality associates include stakeholders, developers, providers, consumers, QoS brokers, quality assurers, and quality managers. In the case of quality activity, a number of activities necessary in web service delivery comprise the measures of quality. Such activities include contracts, clarification, search, delegation, development, registration, report, notification, monitoring, and management that aggregately cover the different aspects of web service delivery. (Kim Lee 2005)

Van Moorsel (2001) explained that metrics, especially quantitative metrics, are useful in applying quality of service. Quantitative metrics are able to evaluate quality of web services by quantifying web service factors, monitoring changes in quantifiable web

service data, and assesses and manages the different online services. As such, quantitative metrics have been developed such as quality of experience or QoE and quality of business or QoBiz. These metrics are able to capture different forms of web services including business to consumer or B2C services, business to business or B2B services, and service utility in relation to the service providers. The importance of quantifiable measures of quality is the ability to track immediate changes and provide approximations of web service quality (Berr & Greiner 2007). However, quantifiable metrics also carry the limitation of not being able to determine reasons for any shifts in quality of web service although it can effectively point out changes and the extent of changes. This means that quantifiable metrics need to be applied with qualitative metrics of web service quality in order to support holistic and long-term improvements in web service quality and the maintenance of web service quality.

Conclusion

Existing literature on quality of service for web services gave rise to the four themes covering the definition, conceptualisation and application of quality of service in the context of web services, importance of quality of service to web services, issues faced in applying quality of service to web service, and measures used in quality of service for web services. Although many literatures discussing the important role of quality of service to web services exist, the commonly used QoS applications have limitations and have to be improved in order to meet the rising web service demand. One way of improving quality of service for web services is using multi-measures that covers not only quantifiable but also qualitative measures and consider different quality perspectives and expectations of stakeholders of web services. Another way of enhancing the role of

quality of service relative to web services is improving QoS software applications to cover emerging needs and demands. This change involves a concurrent updating of the knowledge and skills of the quality associates involved in web service delivery in order to spur innovations on quality of service software for web service providers to maintain and improve the quality of web services and techniques for web users to apply in ensuring quality web services for end users. These areas covering the development and effective practical application of QoS measures to support the determination and monitoring of quality of webs services in real time require further investigation. As such, this supports the conduct of a study on these areas.

Bibliography

Berr, T., & Greiner, U., 2007. *Enhancing quality of Web-Service-Based Cooperative Business Processes – Definition and Monitoring of Quality Constraints*. Saarbrücken, Germany: VDM Verlag Dr. Mueller e.K.

Bielski, L., 2004. The Big Squeeze: Eliminating Data Bottlenecks for Web-Based Applications. *Journal Title: ABA Banking Journal*, 96(5), p.55+.

Bouch, A., Kurchinsky, A., & Bhatti, N., 2000. Quality is in the Eye of the Beholder: Meeting Users' Requirements for Internet Quality of Service. Available at: http://www.hpl.hp.com/techreports/2000/HPL-2000-4.pdf [Accessed 13 January 2007]

Conti, M., Das, S.K., & Shirazi, B.A., 2002. Quality of Service Issues in Internet Web Services. *IEEE Transactions on Computers*, 51(6), p.593-594.

Kim, E., & Lee, Y., 2005. Quality Model for Web Services. Available at: www.oasis open.org/committees/download.php/15442/WSQM-TC-Charter1109.doc [Accessed 13 January 2007]

Lee, K.C., Jeon, J.H., Lee, W.S., Jeong, S.H., & Park, S.W., 2003. QoS for Web Services: Requirements and Possible Approaches, W3C Working Group Note 25 November 2003.Available at: http://www.w3c.or.kr/kroffice/TR/2003/NOTEwsqos-20031125/ [Accessed 13 January 2007]

Li, T.H., 2005. A Hierarchical Framework for Modeling and Forecasting Web Server Workload. *Journal of the American Statistical Association*, 100(471), p.748-763.

Ludwig, H., 2003. Web Services QoS: External SLAs and Internal Policies Or: How do we delivery what we promise?, Proceedings of the Fourth International Conference on Web Information Systems Engineering Workshops. Available at: www.ieeexplore.ieee.org/iel5/9036/28684/01286793.pdf [Accessed 13 January 2007]

Mani, A., & Nagarajan, A., 2002. Understanding quality of service for Web services: Improving the performance of your Web services. Available at: www 106.ibm.com/developerworks/library/ws-quality.html [Accessed 13 January 2007]

Menasce, D.A., 2002. QoS Issues in Web Services. *IEEE Internet Computing*, November December 2002, p.72-75.

Offutt, J., 2002. Quality Attributes of Web Software Applications. *IEEE Software*, vol. 19(2), p.25-32.

O'Sullivan, J., Edmond, D., & Hofstede, A., 2002. What's in a Service? Towards Accurate Description of Non-Functional Service Properties. *Distributed and Parallel Databases*, 12, p.117-133.

Porter, M., 1998. *Competitive Advantage*. New York: Free Press.

Ran, S., 2003. A model for Web Services Discovery with QoS. *ACM SIGecom Exchanges*, 4(1), p.1-10.

Sumra, R., & Arulazi, D., 2003. Quality of Service for Web Services – Demystification, Limitations, and Best Practices. Available at: www.developer.com/services/article.php/2027911 [Accessed 13 January 2008]

Van Moorsel, A., 2001. Metrics for the Internet Age: Quality of Experience and Quality of Business. Available at: http://www.hpl.hp.com/techreports/2001/HPL-2001 179.html [Accessed 13 January 2008]

Wang, H., Huang, J.Z., Qu, Y., & Xie, J., 2004. Web services: problems and future directions. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1, p. 309–320.

Wolter, K., & Van Moorsel, A., 2001. The Relationship between Quality of Service and Business Metrics: Monitoring, Notification and Optimization. Available at: www.hpl.hp.com/techreports/2001/HPL-2001-96.html [Accessed 13 January 2008]