QoS based Cloud Service Provider Selection Framework

Kumar N. and Agarwal S.
Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, UP, INDIA

Abstract

A wide range of service offerings has been opened by the cloud computing industry to the customers. The cloud service providers and their services are increasing exponentially in present scenario of computing. In such a situation a customer faces the problem of selecting the best cloud service provider according to his personalized quality of service requirements. Moreover, the dynamic nature of the underlying network and several unpredictable circumstances, cause the service providers to depart from the promised QoS, as per Service Level Agreement. In this paper the authors present a framework for cloud service selection engine which acts as a tool to enable the customers to select the most appropriate cloud service provider from the Web Repository. The framework uses analytic hierarchy approach for multi-criteria QoS decision making which accelerates the selection process. Past users' experience is used as a heuristic which helps the algorithm to converge in polynomial time. A comparison has been given between proposed mechanism and previous approaches and the results of proposed technique are promising.

Keywords: AHP, cloud service provider, service level agreement, QoS.

Introduction

Cloud service providers provide resources on pay-per-use basis which can be scaled up and down according to specific users' need. This encourages business organizations to host their applications on cloud infrastructures and save large investment or up-front costs. However with a plethora of cloud service providers available on the web spectrum along with a range of functionally equivalent services, it becomes difficult for the customers to select the most suitable, most consistent and most efficient service provider for themselves which delivers best Quality of Service. In this context the issues of ranking cloud service providers and customer satisfaction on the delivered QoS, have received renewed attention in the cloud scenario. Even though the cloud service providers make promises about quality parameters in the Service Level Agreement, most of the times they fail to fulfil them because the internet environment is always subject to unpredictable conditions no matter how robust a system may be. Therefore discovering an optimum service provider for oneself has become a challenge for a customer or an end user.

It is worth noting that in many cases users are not clear regarding which quality attributes are most important for their application. Further, because of the availability of similar service plans with different price structure make the selection problem more complicated. Therefore we model the problem as Multi Criteria Decision Making (MCDM) problem where a number of criteria must be considered before arriving at the final result. For this purpose the QoS criteria must be clearly categorised and classified so that application-centric importance can be decided.

In this direction the Cloud Service Measurement Index Consortium (CSMIC) has laid down certain metrics for evaluation and comparison of the service providers, which are collectively termed as Service Measurement Index [11]. SMI is based on ISO standards and defines seven groups of QoS attributes which act as a foundation on which different providers can be cross compared. The top level groups of the SMI framework include Accountability, Agility, Cost, Performance, Assurance, Security and Privacy, Usability. Within each of these groups lower level attributes are defined. These attributes act as Key Performance indicators of the providers' efficiency. Thus SMI acts as a road map which instigates towards better overall judgement.

The current paper proposes a cloud service selection engine framework which utilises SMI attributes and a MCDM solving technique to rank the available service providers and select the one which satisfies QoS most consistently. The given model is easy to implement and has a simple architecture. The results are obtainable in polynomial time.

Related Work: As cloud computing is becoming one of most popular field of study. The number of individual clients and business organizations prefer to use the cloud services to evade upfront costs. The challenges faced by business enterprises have also been discussed in the literature[1]. Because cloud is a service oriented, pay per use model, there are many aspects of cloud computing viz., resource scheduling, workload balancing, virtual machine migration, etc. which have been studied by the researchers[2]. However all these aspects have one major point of consideration that is, Quality of Service (QoS). A number of

mechanisms have been proposed to increase the QoS of a service provider[3] and much consideration has been given to customer's point of view[4,5]. The idea of giving ranks to cloud service provider is originated[5,6]. The number of quality attributes is large[7]. Expert systems based solutions[8] and statistical solutions[9] are often complex in nature. In this paper we give a simpler solution for declaring the rank of the service providers by formulating the ranking problem as a Multi Criteria Decision Maaking (MCDM) Problem.

The Analytic Hierarchy Process[10,11] is a powerful technique to solve MCDM problems as compared to optimization solutions[12]. AHP has only scarcely been applied in solving cloud computing problems. It can also measure safety metrics[13] in terms of attributes and sub attributes. In this paper cloud model[14] along with AHP is used to rank cloud providers and hence it can be an efficient mechanism towards cloud service selection engine framework. Building quality aware cloud environments[15,16] can therefore be built.

Background

The Cloud Services Measurement Initiative Consortium (CSMIC): The CSMIC has developed a set of business-relevant key performance indicators (KPI) which is termed as Service Measurement Index (SMI). SMI is being adopted by business organizations as a standard method to measure cloud based services. SMI defines seven categories of Quality of Service attributes as shown in figure 1. Within each category there are sub-attributes on the basis of which cloud services can be compared. However, SMI does not specify which attribute is most important for a given application. SMI is a means for understanding and categorizing various IT services floated on the web. The proposed framework for service provider selection uses SMI attributes to rank service providers according to their ability to satisfy specific Quality of Service constraints. It must be noted that QoS has varied level of importance for different individuals. For example a government agency is more interested in security rather than the cost while for a scientific application performance is a more crucial issue. An e-commerce application requires a good combination of all the quality attributes. In this manner SMI acts as a roadmap to understand in clear terms which quality parameters are most important for a particular user.

Analytic Hierarchy Process: The Analytic Hierarchy Process (AHP), is a promising approach for solving Multi Criteria Decision Making (MCDM) problems [2]. In AHP, complex problems are structured into a hierarchy of criteria, sub-criteria and decision alternatives from which the final choice is to be made. In contrast to a simple MCDM situation the criteria in AHP may be expressed in different units eg, time, currency, CO_2 emission, etc and are often conflicting in nature. It must be noted over here that QoS criteria in Clouds have several dimensions and it is due to this reason AHP offers a great assistance in selecting an optimal service provider or the best service from among the given alternatives. To make decisions using AHP one identifies and analyses the tradeoffs between different alternatives to achieve an objective. A number of paired comparisons between the criteria and the alternatives are done iteratively. The final result is an ordered list alternatives according to user's preference. The structure of a typical AHP problem and its solution can be read from [4]. The most important task in AHP is to perform several pair-wise comparisons by using a scale. Such a scale is a mapping between the textual descriptions of a criterion to its quantitative measure. The scale given by Satty is given in table 1.

Service Request: The requests for various types of resources as service originate from heterogeneous points all over the world. For the sake of simplicity and understanding the model shows how a single request is being handled by the search engine framework. Users have personalized QoS constraints which they submit to the system. Not all the cloud providers may qualify for satisfying users' needs therefore further components in the system will only extract those candidates that are best suited for the user with the given quality requirements.

Top Level Groups	Sub-Attributes
Accountability	Auditibility, Compliance, Contracting Experience, Data Ownership, Ease of doing business, Provider Usability, Stability, Provider Certification, Provider Ethicality, Provider Personnel Requirements, Provider Supply Chain, Security Capabilities, Sustainability
Agility	Adaptability, Capacity, Elasticity, Exensibility, Flexibility, Portability, Scalability
Assurance	Availability, Maintainability, Recoverability, Reliability, Resiliency, Service Stability, Serviceability
Financial	Acquisition and Training Cost, On-going Cost, Profit Sharing
Performance	Accuracy, Functionality, Suitability, Interoperability, Service Response Time
Security and Privacy	Access Control, Political, Data Integrity, Data Privacy, Data Loss, Physical and Environmental Security, Threat Management, Retention
Usability	Accessability, Client Personnel Requirements, Installability, Learnability, Operability, Suitability, Transparency, Understandability

Figure-1
Categories of QoS Attributes

Table- 1
Pair wise Comparison Scale

| Relative Importance	Definition	Description
1	Of same Importance	Two criteria are of equal importance
3	Essential or strong importance	Experience strongly favors one criteria over another
7	Demonstrated importance	A criteria is strongly favored and its dominance is demonstrated in practice.
9	Absolute importance	The evidence favoring one criteria over another is of highest possible order of affirmation
2,4,6,8	Intermediate values between the two adjacent judgements	Whenever a compromise is needed
Reciprocals of above non-zero values	If criteria i has one of the above non-zero value assigned to it when compared with criteria j, then j has a reciprocal value when compared with i	
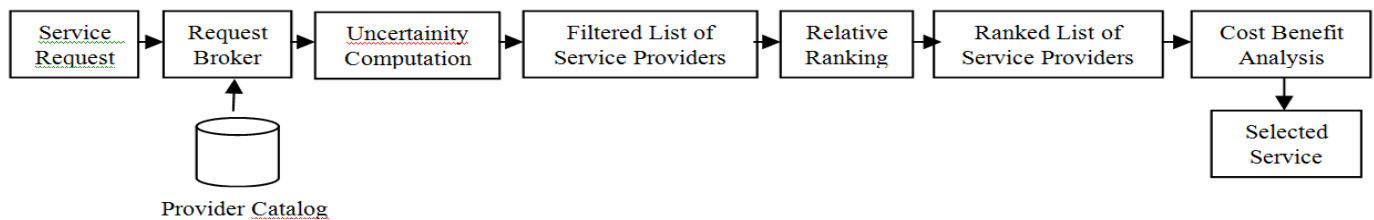
Figure-2
The Proposed System Model

System Model: In this section we show the main components of the proposed system model of QoS based Cloud Service Provider Selection Framework.

Request Broker: This component acts as an interface between the user and the selection framework. It interprets the quality requirements of the users' application and assigns relative weights to the QoS parameters submitted by the user. The request broker has access to cloud resource provider catalog with the help of which it produces a list of candidates that will be able to satisfy users' requirements.

Uncertainty Computation: This component reduces the search space by filtering and producing the most consistent resource providers from the available list. This process saves computation time and helps to reach at the final choice quickly. The main idea behind uncertainty computation is to measure by how much the delivered quality values depart from the promised quality levels as per the Service Level Agreement. For this purpose the past entries of the transaction log are scanned and the variance is measured for a given QoS criteria. The cloud model given by [8] has been used for this purpose. We explain the idea with the help of an example.

Let CSP_1 and CSP_2 are the two service providers which are to be judged for consistent performance on QoS criteria C_1. C_1 can be a quality parameter like Response Time, Turnaround Time and Availability etc. The performances of CSP_1 and CSP_2 are

recorded by a series of transaction log entries which reflect the actual C_1 values delivered (and hence the fluctuations) by each provider during real time invocation. We illustrate five such log entries of CSP_1 and CSP_2 in the table 2.

Table-2
Uncertainty Computation

Log entries of actual QoS Values			
CSP$_1$		CSP$_2$	
Transaction No.	C$_1$ Value	Transaction No.	C$_1$ Value
CSP$_1$(1)	39	CSP$_2$(1)	26
CSP$_1$(2)	36	CSP$_2$(2)	50
CSP$_1$(3)	40	CSP$_2$(3)	41
CSP$_1$(4)	36	CSP$_2$(4)	44
CSP$_1$(5)	36	CSP$_2$(5)	22
Average Value	37.4	Average Value	36.6

From table 2 it is clear that the Average QoS Value of C_1 given by CSP_1 is larger than that of CSP_2. CSP_2 will obviously be preferred over CSP_1. However further analysis of individual transaction reveals that, in three transactions CSP_1 has provided a lower value of C_1 than CSP_2. This means that CSP_2 has more fluctuations on C_1 and therefore CSP_1 proves out to be more consistent than CSP_2.

Filtered List of Service Providers: In this step we obtain the list of resource providers which have proven consistency in their

quality level. The filtered list provides a reduced search space from where the final selection has to be done. This list is updated periodically to introduce new resource providers that started showing consistency in the recent past and to eliminate those candidates that showed a clear drop in their performance.

Relative Ranking: At this point we have already shortlisted the consistent resource providers. Now, our aim is to provide an ordering or rank to them so that they can be compared on a number of quality parameters. Ranking the resource providers can be viewed as Multi Criteria Decision Making (MCDM) problem which has been discussed extensively in the literature. One of the most widely accepted solution of MCDM problem by the researchers is the Analytic Hierarchy Process as mentioned in section III. Suppose that there are 'n' number of service providers and 'm' QoS criteria, we apply AHP to compare each of the 'n' service providers on each of 'm' criteria and also to evaluate the importance of 'm' QoS criteria itself. The final result is priority ranking of the service providers which is passed on to the next phase for further analysis.

Ranked List of Service Providers: In this phase we obtain a ranked list of service providers which are most suitable for the user's need. The ranking is given on the basis of a score as obtained after the application of AHP. The top ten cloud service providers are identified and the list is passed on to the next phase for further processing.

Cost Benefit Analysis: This is an evaluation phase in which the selected service providers are analysed on the basis of the benefits obtained as against the cost that they charge. The analysis is presented to the user in the form of a report. The user can finally select the service provider according to his quality preferences and budget constraints.

Selected Services: This is the final phase where the service and the service provider have been selected and presented before the user.

Provider Catalog: The provider catalog serves as a repository where globally available service providers and their quality information are stored. The past users' experience of these service providers is also recorded in this repository. The data repository is consulted by the request broker to decide at the preliminary stage, the available and suitable service provider for the current request. The repository is updated from time to time to reflect the changing performance of the service providers as well as other related dynamic information such as cost structure, availability, etc.

Experiment Results and Discussion

For the experiments we created our own dataset of transaction log entries for 50 cloud service providers. Using the proposed mechanism we selected top 10 service providers and recorded their performance in a number of ways. Figure 5 shows the comparison among the top ten service providers on the basis of their consistency in performance and the average response time. The cloud service providers 1, 2 and 5 have low average response time and a high consistency level while the cloud service providers 3, 4, 6, 7, 8, 9 and 10 have a large average response time and a low consistency level. Therefore a user may choose from option 1, 2 or 5. The final selection is done on the basis of cost benefit analysis as explained in section III.

The radar graphs shown in figures 6 and 7 represent the evaluation results of the cloud service providers CSP_1 and CSP_2 on five qualities of service parameters. From the graphs it is evident that both CSP_1 and CSP_2 have similar consistency level but cost wise CSP_2 is charging more that CSP_1 for its services. Therefore the final choice made by the user will rely upon the remaining three parameters viz. Response time, throughput and availability. For example, for a real time banking application CSP_1 will be a better choice as availability is high in this case. For a scientific application where accuracy of results is of more importance, CSP_2 will be a better choice.
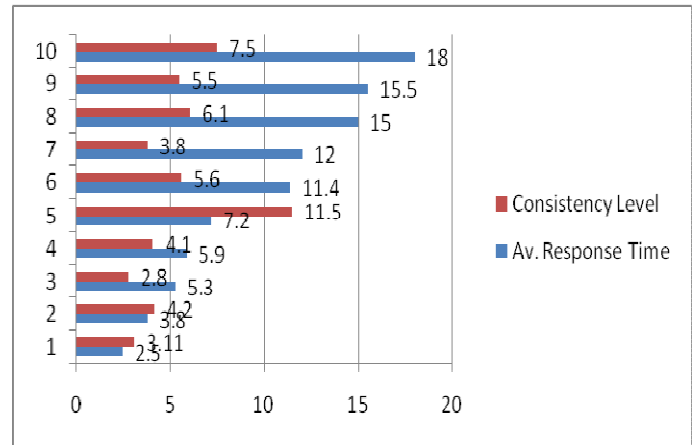

Figure-3
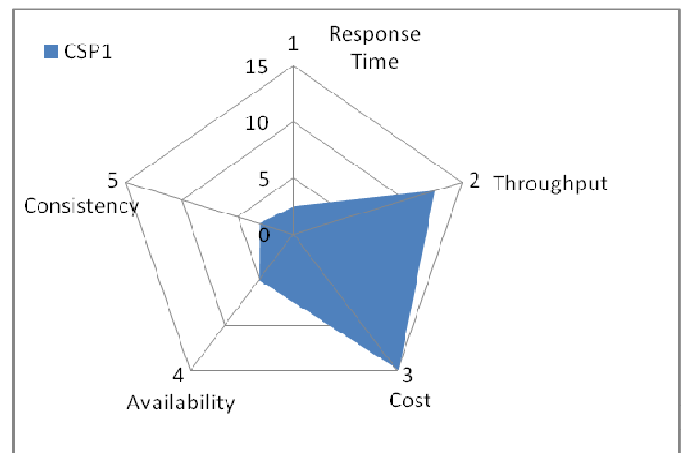**Comparison of top ten service providers to show their consistency level**

Figure-4
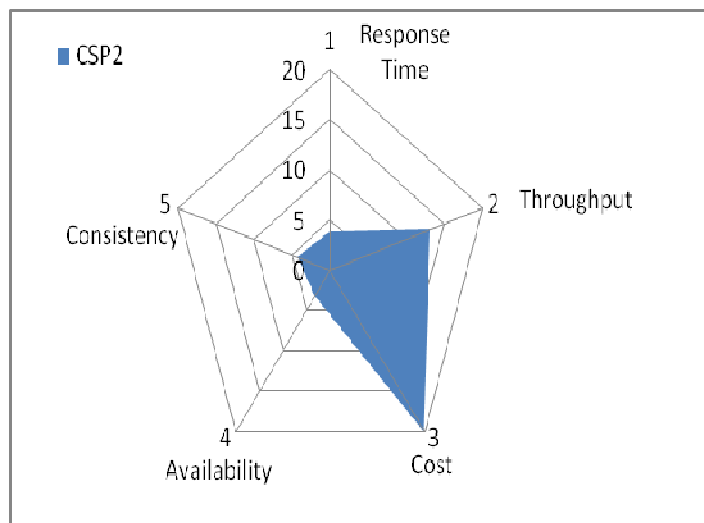**Evaluation of CSP_1 on five quality parameters**

Figure-5
Evaluation of CSP$_2$ on five quality parameters

Conclusion

In this paper the authors have a proposed a service provider selection model which can be viewed as the underlying framework for Cloud Service Selection Engine. The relevance of the model is high in the present scenario because during the last four years Cloud has emerged as the most promising service oriented computing paradigm and the number of cloud service providers along with their varied service plans and schemes have grown exponentially. As a result, end user or customer faces the problem of selecting best cloud provider for oneself. The model tackles this problem by evaluating available cloud service providers on a number of quality parameters and then provides a ranking to top ten candidates. The strong aspect of the proposed model which makes it different from the earlier work is that we also perform consistency analysis of the selected service providers. Experiments show that the implementation of the model is simple and the results are obtained in polynomial time. The user can easily choose the best option depending on the personalized needs.

For future work, it is planned to implement this model on a real world data set and evaluate the results using a simulation environment. Also other techniques of Multi Criteria Decision Making problems will be applied and comparison with the present work will be done.

References

1. Buyya R., Garg S.K. and Calheiros R.N., SLA Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture and Solutions, Proceedings of the International Conference on Cloud and Service Computing, IEEE, Australia, 1-10, **(2011)**

2. Beloglazov A. and Buyya R., Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud DataCenters Under Quality of Service Constraints, *IEEE Transactions on Parallel and Distributed Systems*, **24(17),** 1366-1379 **(2013)**

3. Zheng Z., Wu Z., Zhang Y., Lyu M.R. and Wang J., QoS Ranking Prediction for Cloud Services, *IEEE Transactions on Parallel and Distributed Systems,* **24(6),** 1213-1222 **(2013)**

4. Ding S., Yang S., Zhang Y., Liang C. and Xia C., Combining QoS prediction and customer satisfaction estimation to solve cloud service trustworthiness evaluation problems, *Knowledge Based System,* **56,** 216-225 **(2014)**

5. Zhao L., Ren Y., Li M. and Sakurai K., Flexible service selection with user-specific QoS support in service-oriented architecture, *J. Network and Computer Applications,* **35(3),** 962-973 **(2012)**

6. Garg S.K., Versteeg S. and Buyya R., A framework for ranking of cloud computing services, *Future Generation Computer Systems,* **29(4),** 1012-1023 **(2013)**

7. Cloud Service Measurement Index Consortium (CSMIC), SMI framework. URL:http://beta-www.cloudcommons.com/servicemeasurementindex **(2014)**

8. Mahmoodi K.R., Nejad S.S. and Ershadi M., Expert Systems and Artificial Intelligence Capabilities Empower Strategic Decisions: A Case study, *Res. J. Recent Sci.,* **3(1),** 116-121 **(2014)**

9. Movahedi M.M., A Statistical Method for Designing and analyzing tolerances of Unidentified Distributions, *Res. J. Recent Sci.,* **2(11)**, 55-64 **(2013)**

10. Saaty T.L., Decision making, new information, ranking and structure, *Mathematical Modelling,* **8,** 125–132 **(1987)**

11. Triantaphyllou E., Stuart H. and Mann S.H., Using the analytic hierarchy process for decision making in engineering applications: some challenges, International Journal of Industrial Engineering: Applications and Practice, **2(1),** 35-44 **(1995)**

12. Sanjay J. and Nitin A., An Inverse Optimization Model for Linear Fractional Programming, *Res. J. Recent Sci.,* **2(4),** 56-58 **(2013)**

13. Khan S. and Al Ajmi F.M., Cloud_Computing_Safety Concerns_in_Infrastructure_as_a_Service, *Res. J. Recent Sci.,* **3(6),** 116-123 **(2014)**

14. Wang S., Zheng Z., Sun O., Zou H. and Yang F., 'Cloud model for service selection', *IEEE Conference on Computer Communications Workshops* (INFOCOM WKSHPS), 666 -671 **(2011)**

15. Kourtesis D., Alvarez-Rodríguez J.M. and Paraskakis I., Semantic-based QoS management in cloud systems: Current status and future challenges, *Future Generation Computer Systems*, **32,** 307-323 **(2014)**

16. Salam A., Nadeem A., Ahsan K., Sarim M. and Rizwan K.., 'A_class_based_QoS_model_for_Wireless_Body_Area Sensor_Networks, *Res. J. Recent Sci.,* **3(7),** 69-78 **(2014)**

17. Kafetzakis E., Koumaras H., Kourtis M.A. and Koumaras V., QoE4CLOUD: A QoE-driven multidimensional framework for cloud environments, International Conference on Telecommunications and Multimedia (TEMU), 77-82 **(2012)**