

A Review on Workload Prediction of Cloud Services

Supreet Kaur Sahi
Research Scholar
Jaipur National University, Jaipur

V.S.Dhaka
Professor and Head Dept. Of CSE
Jaipur National University, Jaipur

ABSTRACT

In cloud computing scalable and IT-enabled capabilities are delivered as a service to consumers using Internet. Workload prediction for cloud services differs from one application to another. In this paper review is presented about different work done for workload prediction of cloud computing for e-business websites. Findings and future scope are also discussed in this paper.

Keywords

E-Business, SaaS, PaaS, IaaS.

1. INTRODUCTION

Workload prediction plays very important role for any company. According to Kohavi [1] with every 100 ms increase in load time for Amazon.com sales are decreased by 1%. According to Zona Research, report published in April 1999, US have lost \$43.5 billion in the 1998 in e-commerce as a result of bad performance [2]. According to its statistics, in the holiday season of 1998, over 1/3 of customers gave up due to slowness and 44 percent had used conventional stores. The term workload with respect to cloud computing is an extent to which cloud service consumers put their virtual systems on the cloud environment. The way IT services are delivered to different customers are changed by cloud computing. The key attribute of cloud computing are as follow [3]:

- 1) On-Demand Resource Allocation: - The resources are allocated when required and will be billed accordingly.
- 2) Scaling: - The user is given illusion of infinite number of resources.
- 3) Multitenant environment: - Consumers are provided with resources from single cloud implementation hence saving the cost.

The four layers of cloud are as follow:

- 1) Software as a Service (SaaS): - In this layer applications are delivered on subscription.
- 2) Platform as a Service (PaaS): - It consists of software development framework delivered and it is charged on usage basis.
- 3) Infrastructure as a Service: - In this layer network, storage and computation facilities are delivered to users as used by them.
- 4) IT foundation provides basic building block for using above layers.

Cloud computing results in major workload shift. Major workload is handled by cloud. The user side hardware and software requirements are decreased. Different applications on cloud have different workload. As for example workload of some e-business website will be higher during festive season. Online financial services require high amount of security whereas scientific computing on cloud require high computing services.

In this paper the review of different work done for workload prediction of cloud based e-business websites is presented. Further the findings and future scope is discussed.

2. BACKGROUND

Menasce [2] discusses different challenges in scaling e-business sites when workload changes. The author present a multi layer reference model for analysis and capacity planning of e-business websites. The model takes customer behavioral pattern into account for explaining fluctuating workload. The author take into account four layers base on business model, functional model, customer model and resource model. The author suggests that workload characterization must be done at all these levels.

Almedia [4] discusses about different steps involved in capacity planning of web services. Different steps as discussed by author are as follow: -

- 1) Understanding the environment: -It includes understanding of different features like servers, connectivity, protocols, policies etc.
- 2) Workload characterization: - In this author discusses workload of different components of website. The session detail can be used for finding workload. Different parameters for workload model can be number of users visited, peak traffic time, shopper characteristics, demand of resources, etc.
- 3) Performance model development and validation: - Performance model will be based on different input parameters and validations will be based on difference between target and actual system.
- 4) Analysis of future scenario: - The final step is to analyze future scenario in order to have cost-effective plan.

Raquel [5] proposes a model for capacity planning activity of cloud-based infrastructure. The author has focused on IaaS. The author acquires resources from different market, which are on-demand, reservation and on-spot market. In this paper author prepare an analytical model for dynamic cloud infrastructure usage plan. This model takes different scenarios for evaluation. Finally the model is able to find out more profitable plan by finding the number of instances required by application on cloud. Wang [6] discusses pros and cons of ecommerce development environment. Bhaskar [7] provides comparison of different solution procedures on the basis of architecture, virtualization, servicing, balancing, security etc.

3. FINDINGS

By reviewing above discussed literature it has been found that all work related to workload prediction is theoretical and analytical. The author has found that Menasce [2] has done workload optimization at different levels of multi-layered model by using concepts of clustering and session details. In [2] mathematical metrics for finding workload of different sessions was suggested. By doing workload prediction at each level of Figure 1 more accurate results can be generated.

The author has also used customer behavior graph for capacity planning. Different steps of capacity planning are very well explained by Almedia [4]. Figure 2 shows different

steps involved in capacity planning as presented by Almedia [4].

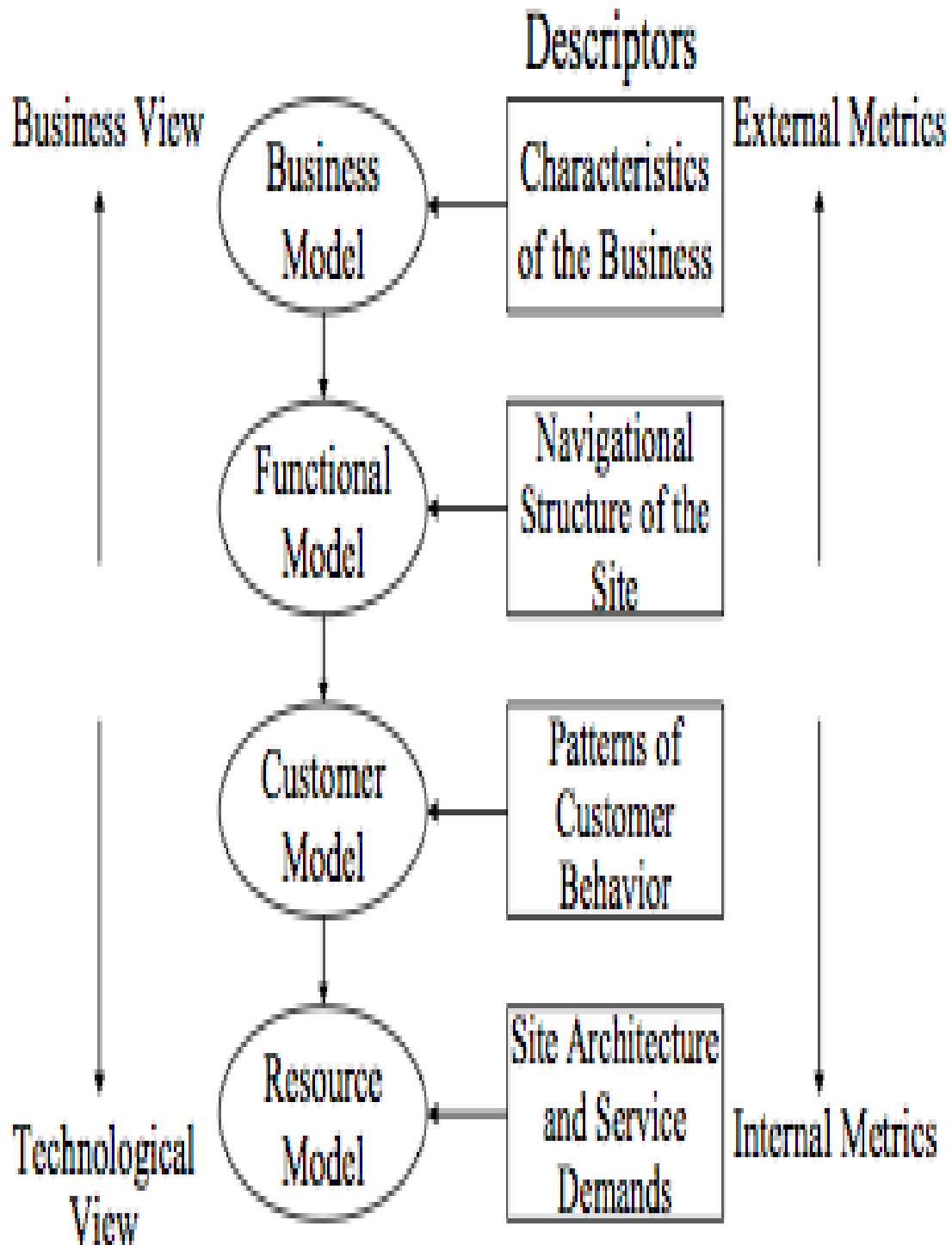


Figure 1: A multi-layered Model for E-Business sites [2].

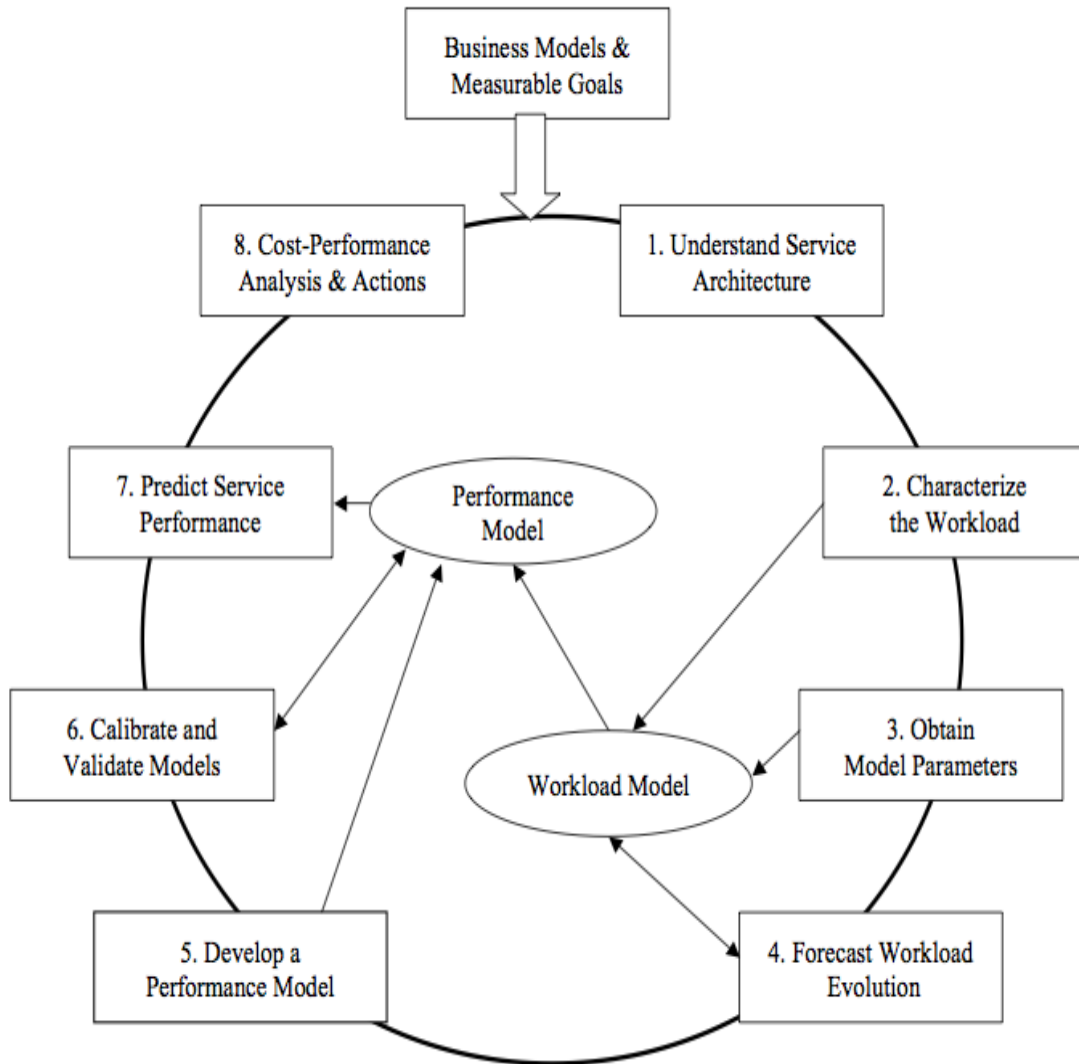


Figure 2: Steps involved in Capacity Planning [4]

This paper gives steps for analyzing web services. This paper can be used as a base for workload prediction for different applications of cloud services. Raquel [5] uses steps of Almedia [4] as a base for capacity planning of cloud based infrastructure for web applications. Analytical model was also proposed for using cloud infrastructure in cost effective manner. In this model different parameters like mean arrival, service rate, peak time, utility gain, penalty are taken into account. Finally a mathematical formula was derived taking into account all this parameters. The author has used this model for maximizing profit from web applications. By reviewing Bhaskar [7] the author is able to compare different cloud systems like Amazon EC2, Sun Cloud, Eucalyptus, Google App Engine, Azure and RightScale. It has been found that in Amazon EC2 applications of users are controlled as instances not as services. Eucalyptus is an open-source implementation of Amazon Web Services (AWS) clouds system. Sun Cloud is agile in nature and provides multiple architectures of underlying hardware. Google manages Google App Engine whereas Microsoft has hosted Azure. They both are PaaS and SaaS providers. RightScale provides an efficient failover method. Following table summarized different findings by an author:

Table 1: Findings

Sno	Paper Studied	Findings
1.	Menasce [2]	<ul style="list-style-type: none"> ➤ Multi Layered model discussed. ➤ The author has used customer behavioral pattern for website's workload characterization. ➤ Customer Behavior Model Graphs are used for calculating different metrics in order to find workload.
2.	Almedia [4]	<ul style="list-style-type: none"> ➤ Different steps involved in capacity planning are discussed. ➤ This paper provides base for different activity of workload prediction.
3.	Raquel	<ul style="list-style-type: none"> ➤ Analytical model is proposed

	[5]	<p>in this model.</p> <ul style="list-style-type: none"> ➤ IaaS model taken into account in this paper. ➤ Author takes into account different parameters. ➤ The paper proposed the way to acquire resources from cloud service providers in order to increase profitability.
4.	Wang [6]	<ul style="list-style-type: none"> ➤ The paper provides different parameters like security, technology, cost, mobile terminal etc. for implementation of e-commerce development model on cloud technology.
5.	Bhaskar [7]	<ul style="list-style-type: none"> ➤ The author compare different cloud system on the basis of architecture, virtualization, storage, load balancing, interoperability, programming framework, security etc. EUCALYPTUS is an open source of AWS. The SUN CLOUD is an agile in nature. Amazon Web Services has Amazon programming framework whereas Flexiscale support C, C++, Java, and Perl.

4. CONCLUSIONS AND FUTURE SCOPE

In cloud computing user pay only for the amount of services used. There are number of cloud service providers which can be selected on the basis of application needs. The models available for workload prediction of cloud environment till date are analytical or mathematical. The steps involved in capacity planning activities are available. They can be taken as base for workload prediction of different applications. Work can be done in future for automation of workload prediction of different cloud services. Different forecasting tools like Matlab, MS Excel, S-plus etc. can be used for prediction of workload and forecasting of cloud based services. There are different applications available on the cloud. Workload needs to be application centric as different applications have different requirements.

5. REFERENCES

- [1] Ron Kohavi and Roger Longbotham, Online Experiments: Lessons Learned, IEEE Computer, Vol 40, issue 9, p. 103-105, Sept 2007.
- [2] V Daniel A.Menasce, Virgilio A.F. Almeida "Challenges In Scaling E-Business Sites": Proc. 2000 Computer Measurement Group Conference, Orlando, FL, and December 10-15, 2000.
- [3] "Cisco Cloud Computing - Data Center Strategy, Architecture, and Solutions": 2009 Cisco Systems.
- [4] Virgilio A.F. Almeida "capacity planning for Web Services Techniques and Methodology".
- [5] Raquel Lopes, Francisco Brasileiro, Paulo Ditarso Maciel Jr. "Business-Driven Capacity Planning of a Cloud-based IT Infrastructure for the Execution of Web Applications": 2010 IEEE.
- [6] Xiaofeng Wang "Research on E Commerce development model in cloud computing environment": 2012 International Conference on System Science and Engineering.
- [7] Bhaskar Prasad Rimal, Eunmi choi, Ian Lumb " A Taxonomy and Survey of Cloud Computing Systems": 2009 IEEE Computer Society, DOI 10.1109/NCM.2009.218
- [8] <http://searchcloudcomputing.techtarget.com/feature/What-keeps-Amazon-Web-Services-cloud-on-top>
- [9] <http://aws.amazon.com/what-is-cloud-computing/>
- [10] <http://www.techrepublic.com/blog/the-enterprise-cloud/top-cloud-iaas-providers-compared/>
- [11] <http://finance.yahoo.com/news/research-markets-global-cloud-computing-133200619.html>
- [12] <http://www.forbes.com/sites/louiscolombus/2012/11/08/cloud-computing-and-enterprise-software-forecast-update-2012/>
- [13] <http://aws.amazon.com/ec2/#pricing>
- [14] <http://www.peer1.com/knowledgebase/how-slow-website-impacts-your-visitors-and-sales>
- [15] <http://www.ibmpressbooks.com/articles/article.asp?p=1927741>