

# Workload Prediction of Cloud Services

Dr. Manjot Kaur Bhatia, Urvashi Sapra, Riya Bhatia, Pratibha Chaudhary

*Professor, JIMS*

*Student, JIMS*

*Student, JIMS*

*Student, JIMS*

Date of Submission: 08-12-2022

Date of Acceptance: 16-12-2022

## ABSTRACT

Cloud Computing refers to the servers that are accessed over the internet (present at remote location), In simple terms, it means storing, managing, and accessing the data & programs on the remote servers that are hosted on internet instead of computer's harddrive. In short we store, manage & process data on remote servers. The major variable is workload prediction which is used to measure the efficiency and operational cost that can help cloud to improve. Accuracy is the major component in prediction of workload. Cloud computing had seen a long journey from the start till today.

## I. INTRODUCTION

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 stated statistics, the season of vacations of 1998, more than 1/3 customer left due to slowness and about 44 percent 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 key attribute of cloud computing are as follow [3]:

1. On - demand Self services: The Cloud computing services does not require any human administrators, user can monitor and manage computing resources whenever needed.
2. Broad network access: The services regarding computing is only provided to standard networks and heterogeneous devices.
3. Rapid elasticity:

The computing services might have IT resources that helps to scale out and in on a faster rate as needed.

## 4. Resource-Pooling:

The IT resource (e.g., networks, servers, storage, applications, and services) present are shared across multiple applications and occupant in an uncommitted manner.

Various service in cloud computing are:

- 1) Infrastructure as a Service (IaaS): means that you rent IT infrastructure from a cloud provider, such as Microsoft Azure or Amazon Web Services
- 2) SaaS (Software as a service): Software as a service (or SaaS) is a way of delivering applications over the Internet—as a service
- 3) Platform as a Service (PaaS): This layer provides the resources to actually build applications.
- 4) IT field provides a basic building block for using above layers.

Cloud computing gives a major workload shift. This load is handled by using cloud computing. The hardware and software at user's end are being decreased. Different workload is being handled by different applications.

For example: some e-business websites will have higher workload during the festive season. Online financial services require high amount of security and also require high computing services.

In this paper the different workload prediction of cloud based on e-business websites is presented.

## II. BACKGROUND

Menasce [2] discussed different challenges with the purpose of scaling e-business sites when workload changes. The author has presented a multi-layer model for capacity planning and analysis of e-business websites. The model considers customer behavioral pattern into account based on four layer business model, functional model, customer model and resource model. The author had given the characterization of workload that must be done at all these levels. Almedia [4]

gives about different steps involved in capacity planning for web services. The steps can be taken as:-

1. Environment understanding:- this step includes understanding of features like servers, protocols, policies, connectivity, etc.
2. Characterization of workload:- under this it is discussed about different components of websites regarding workload. Various parameter of workload model has been decided like numbers of users visited, peak traffic time, characteristics of shopper, resources demand, etc.
3. Validation and performance model development:- it will be based on different input parameter and also based on various validation on difference between target and actual system.
4. Future scenario analysis:- this is the final step that is to analyze future scenario in order to have effective plan regarding cost.

Raquel [5] developed a model that deals with capacity planning activity of infrastructure which is cloud-based. Raquel also focused on IaaS. The author acquires resources from different market, which are on high demand, on-

spot market and reservations. The author also prepared an analytical model that can be used for dynamic cloud infrastructure usage plan. It goes under various scenarios for evaluation. The model helps to find out profitable plan by giving number of instances required on cloud application. Wang [6] have stated pros and cons of development environment in e-commerce sites. Bhaskar [7] had given different solutions procedures which are generally based on servicing, balancing, security, architecture, virtualization, etc.

### III. REVIEW ANALYSIS

By reviewing above literature, we have found all work which is related to workload is theoretical and analytical. The author found that Menasce [2] have completed workload optimization at various levels including multi-layered model using concepts of session details and clustering. Mathematical metrics for finding workload session was suggested [2]. Workload prediction at every level depicted in figure 1 shows more accurate results that can be generated.

It has also used customer behavior graph for depicting capacity planning. Various steps for capacity planning are very well explained by Almedia [4]. Figure 1 gives different steps included in capacity planning as presented by Almedia [4].

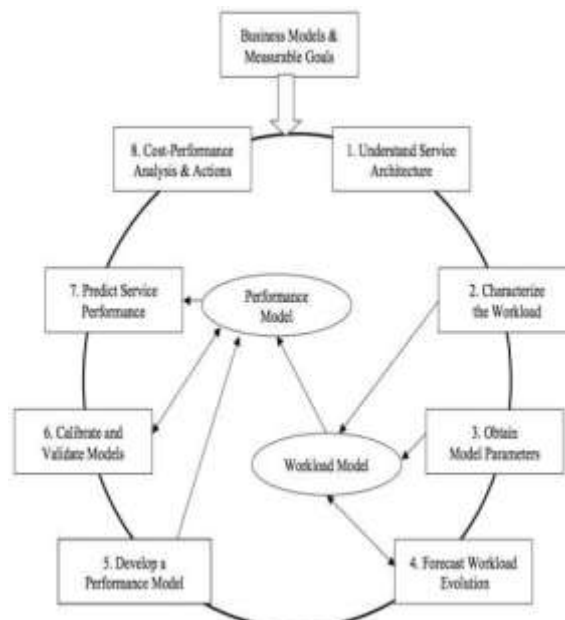


Figure 1. Steps involved in Capacity Planning [4]

The paper gives various steps for web services analysis. This paper acts a base for workload prediction for variety of application associated to cloud services. Raquel [5] used steps of Almedia [4] as source for planning for capacity

for infrastructure which is cloud based associated to web applications. Model related to analysis was also introduced for using cloud infrastructure in cost effective manner. The model contains various parameters like service rate, peak time, mean

arrival, utility gain, penalty is taken into consideration. At the end a mathematical formula was proposed which took all the parameters into account.

The author had used the approach for profit maximizing from web applications. After reviewing Bhaskar [7] author is able to differentiate between various cloud systems like Sun cloud, Amzon EC2, Eucalyptus, Azure and RightScale and Google App Engine. It was discovered that Amazon EC2 users of applications

was controlled not acted as services, where as Eucalyptus is an implementation of open source of Amazon Web Services (AWS) cloud. Sun cloud is based on agile in nature and thus provides multiple types of architectures for underlying hardware. Google is responsible for managing Google App Engine and Microsoft is responsible for hosting Azure. Both includes PaaS and Saas providers. RightScale gives an failover method efficiently. The table summarize some of the findings:

Sno	Paper studied	Finding
1.	Raquel[5]	<ul style="list-style-type: none"> <li>➤ Analytical model was proposed.</li> <li>➤ Major IaaS model was discussed.</li> <li>➤ Various ways to acquire resources from cloud in order to increase profitability.</li> </ul>
2.	Menasce[2]	<ul style="list-style-type: none"> <li>➤ Discussion of multi layered model</li> <li>➤ Usage of customer behavioral pattern for workload prediction.</li> <li>➤ Various customer behavior model graphs used for calculating various metrics to find workload.</li> </ul>
3.	Almedia[4]	<ul style="list-style-type: none"> <li>➤ Capacity planning discussed including its various steps</li> <li>➤ Base for different workload prediction activities.</li> </ul>
4.	Wang [6]	<ul style="list-style-type: none"> <li>➤ The paper provides different parameters like security, technology, cost, mobile terminal etc.</li> </ul>

#### IV. CONCLUSION AND FUTURE SCOPE

In cloud computing all the users need to pay only for the amount of services they use. There are end number of cloud services which can be selected on the basis of application that are needed. All the required model are available for prediction of workload for cloud environment till the date are mathematical or analytical. These steps include capacity planning that have activities into these planning. They can be taken as a base for prediction of workload for various applications. In future work can be done for automation for prediction of workload for cloud services. Different forecasting tools like Matlab, MS Excel, S-plus etc. can be used for prediction of workload and forecasting of cloud services. There are applications that are available on cloud which needs to be centric as different applications have different requirements.

#### 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
- [3] Computer Measurement Group Conference, Orlando, FL, and December 10-15, 2000.
- [4] "Cisco Cloud Computing - Data Center Strategy, Architecture, and Solutions": 2009 Cisco Systems.
- [5] Virgilio A.F. Almeida "capacity planning for Web Services Techniques and Methodology".
- [6] 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.
- [7] Xiaofeng Wang "Research on ECommerce development model in cloud computing environment": 2012 International Conference on System Science and Engineering.
- [8] 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