# PERFORMANCE ANALYSIS OF OPEN SOURCE STORAGE CLOUDS IN CLOUD COMPUTING

Shaveta rani
Student, M.Tech. Deptt. of Computer Sc.
LPU, Jalandhar

Sangeeta sharma
Assistant Professor, Deptt. of Computer Sc.
LPU, Jalandhar

## ABSTRACT

Cloud computing is one of the latest research area that helps in storing the information permanently on the servers and manages the different resources for the requested users to provide on-demand services. In order to create the more usable and economic value based cloud computing, the principles, goals and structure of the cloud engineering is of vital importance. The objective of this study is to analyze the CPU and memory performance of different open source clouds. We will use different open source cloud to measure the different performance metrics like CPU time for downloading and uploading of file, memory usage while downloading and uploading the file, standard deviation of CPU usage and standard deviation of memory usage.

## General Terms

Cloud Computing, Standard deviation, open source clouds etc.

## Keywords

Cloud computing, Metrics, CPU Time, Memory usage, etc.

## 1. INTRODUCTION

One of the major aspects of software engineering is to develop and manage software for commercial use [2]. But now days most organization must continue to use its local resources to operate its daily systems and to store its data but these organization don't use their resource and service efficiently and most of time they remain ideal. This leads to technology where one organization shares its service, resource and store its data on another organization server's known as cloud technique [1] [6]. So, whenever capacity or service is required at short notice it can be accessed by using linked server capacity by paying some considerable fee for such facilities. Here comes the need and role of cloud engineering. Cloud engineering is the study and applied research of the application of the engineering to "cloud".

Cloud computing [3] [9] is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing storage, memory, processing and bandwidth.

Clouds [1] can be defined as large pool of data storage that can be easily used and access virtualized resources such as hardware, development platform and services. In simple words one can say that cloud is the web of different types of server which share resources.

There are numbers of different cloud storage [3] [4] systems, some of these have a very specific focal point, such as storing Web e-mail messages or digital pictures. Others are available to store all forms of digital data. Some cloud storage systems are small operations, while others are so large that the physical equipment can fill up an entire warehouse. The facilities that keep cloud storage systems are called **data centers**.

In simple terms a cloud storage system needs just one data server connected to the Internet. In this case client sends replica of files over the Internet to the data server, which then records the information. When the client wishes to retrieve the information, it accesses the data server through a Web-based interface. The server then either sends the files back to the client or allows the client to access and manipulate the files on the server itself.

Cloud storage [4] systems generally rely on hundreds of data server. Because computers occasionally require maintenance or repair, it's important to store the same information on multiple machines. This is called **redundancy**. Without redundancy, a cloud storage system couldn't ensure clients that they could access their information at any given time. Most systems store the same data on servers that use different power supplies. That way, clients can access their data even if one power supply fails.

Not all cloud storage clients are worried about running out of storage space. They use cloud storage as a way to create backups of data. If something happens to the client's computer system, the data survives off-site. It's a digital-age variation of "don't put all your eggs in one basket."

## 2. PROBLEM DEFINTION

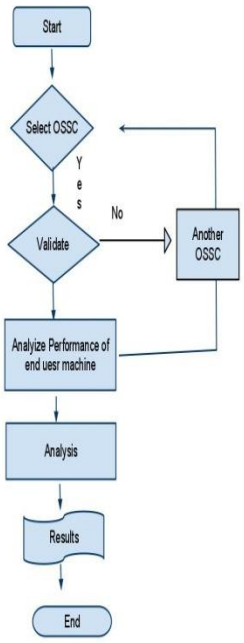
With the passage of time open source storage clouds [7][8] are getting attention. There are various different open source storage clouds exist such as Amazon cloud, google docs, Adrive, open drive etc. Cloud is one of the major challenges being faced by the SaaS users, because the introduction of cloud computing changes our thinking as what is considered to be "our system" and "our data" is no longer physically stored on a specific set of computers and disks, but rather both the concept of system and the locus of our data have evolved into something diffuse geographically distributed. In a cloud data is totally isolated, user does not know where the data is stored so there may be more effect on the performance of the end user machine. In this paper especially I am going to concentrate on analysis of the performance of end user machine in the virtual environment of open source storage clouds, such as google docs, dropbox, and skydrive. Different experiment will be conducted to measure the performance of end machines using different storage clouds. The above mentioned analysis will be conducted by following considering following performance tests on various Open source storage clouds.

Memory stress test : RAM utilization based Tests will be conducted to check the stress on system memory in different OSSCs environment.

CPU test: To test the effects of processor interference due to background load if any, it should be analyzed how each instance uses the processor.

## 3. ANALYSIS

Software metrics one of the important concepts of software engineering plays a great role in coordinating [5] and managing the software project. With the help of software metrics one is able to compute and analyze various attributes of a software project. I have analyzed different open source storage clouds while uploading and downloading the files. We have focused on computing different performance metrics like CPU time, System, Memory Usage, and their variation over different storage clouds.




The following algorithm is used for analyzing the different open source storage clouds.

Step 1: Read Z

Step 2: For I=1 to Z

Step 3: Read Filesize F[I]

Step 4: Read units U of file

Step5: Now convert unit U of the Filesize F[I] in Bytes called UB [i]

Step 6: Read data transmission speed

Step 7: Read data transmission unit DU

Step 8: Convert DU into bytes called DTSB (Data Transmission Speed in Bytes)

Step 9: Download time DT[i]=UB/ (DTSB*60) minutes

Step 10: Read memory usage MU[I] while uploading/downloading a file

Step 11: set X[i]=DT[i], compute square of X[i]

Step 12: Set Y[i]=MU[I], compute Square of Y[i]

Step 13: Set I=I+1

Step 14: End For

Step 15: Now Computer mean of X by using the formula XAM=∑x(i)/Z

Step 16: Similarly compute the mean of Y by using the formula YAM=∑Y(i)/Z

Step 17: Compute Standard deviation of X (CPU Time) by using formula $\sigma = \sqrt{\frac{\varepsilon x^2}{Z} - \left(\frac{\varepsilon x}{Z}\right)^2}$

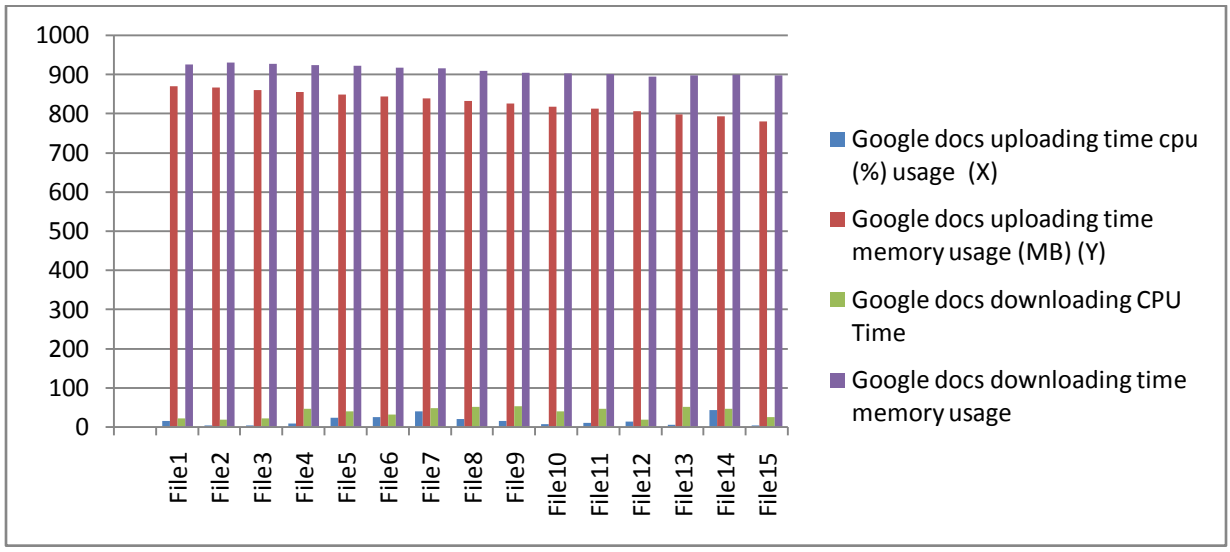
Step 18: Compute standard deviation of Y (memory usage) as $\sigma = \sqrt{\frac{\varepsilon Y^2}{Z} - \left(\frac{\varepsilon Y}{Z}\right)^2}$

Case I: The following table shows how memory and CPU performs during uploading and downloading for selected set of files.

| File name | Google docs uploading time CPU (%) usage (X) | Google docs uploading time memory usage (MB) (Y) | X2 | Y2 | Google docs downloading Time CPU usage | Google docs downloading time memory usage | X2 | Y2 |
|---|---|---|---|---|---|---|---|---|
| File1 | 14 | 869 | 196 | 755161 | 21 | 925 | 441 | 855625 |
| File2 | 4 | 866 | 16 | 749956 | 18 | 929 | 324 | 863041 |
| File3 | 4 | 859 | 16 | 737881 | 21 | 926 | 441 | 857476 |
| File4 | 8 | 854 | 64 | 729316 | 45 | 924 | 2025 | 853776 |
| File5 | 23 | 848 | 529 | 719104 | 40 | 922 | 1600 | 850084 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| File6 | 25 | 843 | 625 | 710649 | 31 | 916 | 961 | 839056 |
| File7 | 39 | 838 | 1521 | 702244 | 47 | 915 | 2209 | 837225 |
| File8 | 20 | 831 | 400 | 690561 | 51 | 909 | 2601 | 826281 |
| File9 | 14 | 825 | 196 | 680625 | 53 | 903 | 2809 | 815409 |
| File10 | 7 | 817 | 49 | 667489 | 40 | 902 | 1600 | 813604 |
| File11 | 10 | 812 | 100 | 659344 | 45 | 901 | 2025 | 811801 |
| File12 | 13 | 806 | 169 | 649636 | 18 | 894 | 324 | 799236 |
| File13 | 5 | 798 | 25 | 636804 | 50 | 897 | 2500 | 804609 |
| File14 | 42 | 793 | 1764 | 628849 | 45 | 899 | 2025 | 808201 |
| File15 | 4 | 779 | 16 | 606841 | 25 | 897 | 625 | 804609 |

The following chart shows the graphical representation of CPU and memory performance of Google docs cloud storage for uploading and downloading the files.



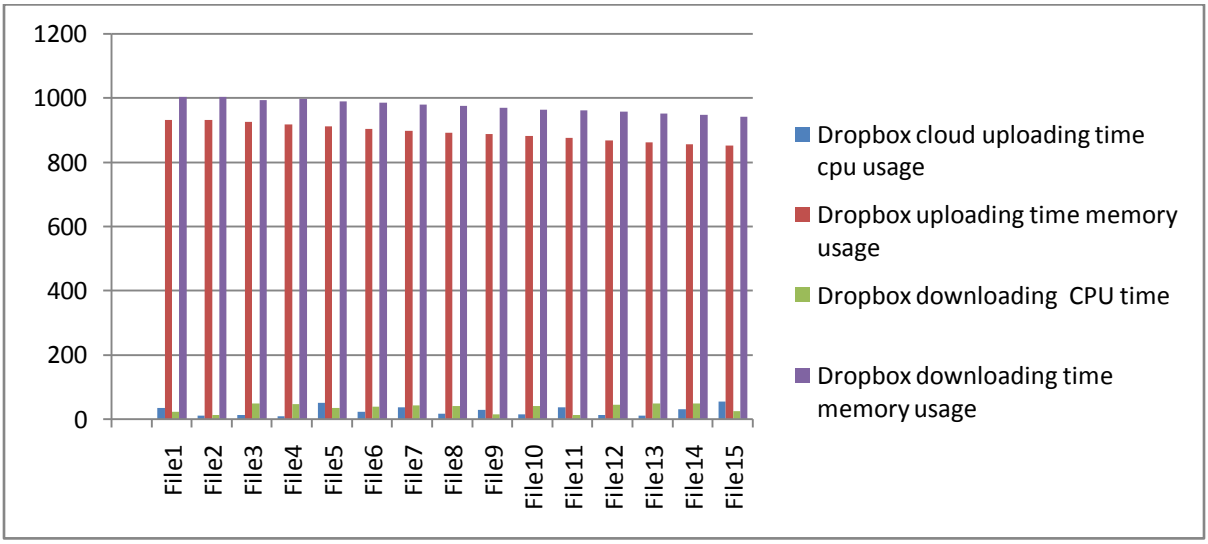

From the above representation it is clear that uploading CPU time on Google doc is lesser than the downloading CPU time. And in regard to memory usage downloading file consumes more system memory as compare to uploading files.

Case II: the following table shows how CPU and memory performs on dropbox cloud for the same set of files.

| File name | Dropbox cloud uploading time CPU usage | Dropbox uploading time memory usage | X1 | Y1 | Dropbox downloading Time CPU usage | Dropbox downloading time memory usage | X2 | Y2 |
|---|---|---|---|---|---|---|---|---|
| File1 | 36 | 932 | 1296 | 868624 | 24 | 1003.52 | 576 | 1007052 |
| File2 | 12 | 931 | 144 | 866761 | 13 | 1003.52 | 169 | 1007052 |
| File3 | 13 | 925 | 169 | 855625 | 50 | 993.28 | 2500 | 986605.2 |
| File4 | 9 | 917 | 81 | 840889 | 47 | 997 | 2209 | 994009 |
| File5 | 52 | 911 | 2704 | 829921 | 35 | 990 | 1225 | 980100 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| File6 | 23 | 904 | 529 | 817216 | 40 | 985 | 1600 | 970225 |
| File7 | 38 | 898 | 1444 | 806404 | 43 | 980 | 1849 | 960400 |
| File8 | 17 | 892 | 289 | 795664 | 41 | 975 | 1681 | 950625 |
| File9 | 30 | 887 | 900 | 786769 | 16 | 970 | 256 | 940900 |
| File10 | 15 | 881 | 225 | 776161 | 41 | 964 | 1681 | 929296 |
| File11 | 38 | 875 | 1444 | 765625 | 14 | 961 | 196 | 923521 |
| File12 | 14 | 868 | 196 | 753424 | 45 | 958 | 2025 | 917764 |
| File13 | 11 | 862 | 121 | 743044 | 49 | 951 | 2401 | 904401 |
| File14 | 31 | 856 | 961 | 732736 | 49 | 947 | 2401 | 896809 |
| File15 | 55 | 851 | 3025 | 724201 | 25 | 942 | 625 | 887364 |

the following chart shows the graphical representation of CPU performance and system memory usage while uploading and downloading files over dropbox cloud storage.



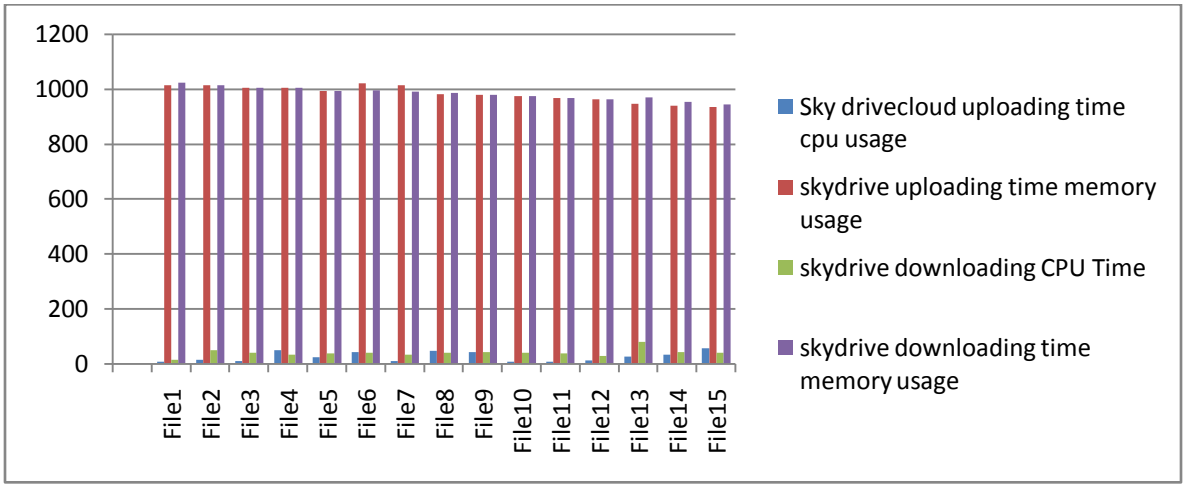

From the above graphical representation it is clear that the relationship between CPU uploading and CPU downloading time is not as clear as in Googledoc, because in this sometimes the CPU uploading time is more and sometimes CPU downloading time is more. However like Google docs the memory performance is same that is downloading file consume more system memory as compare to uploading file in dropbox cloud.

Case III: The following table shows the performance of CPU and memory for the set of files over skydrive cloud.

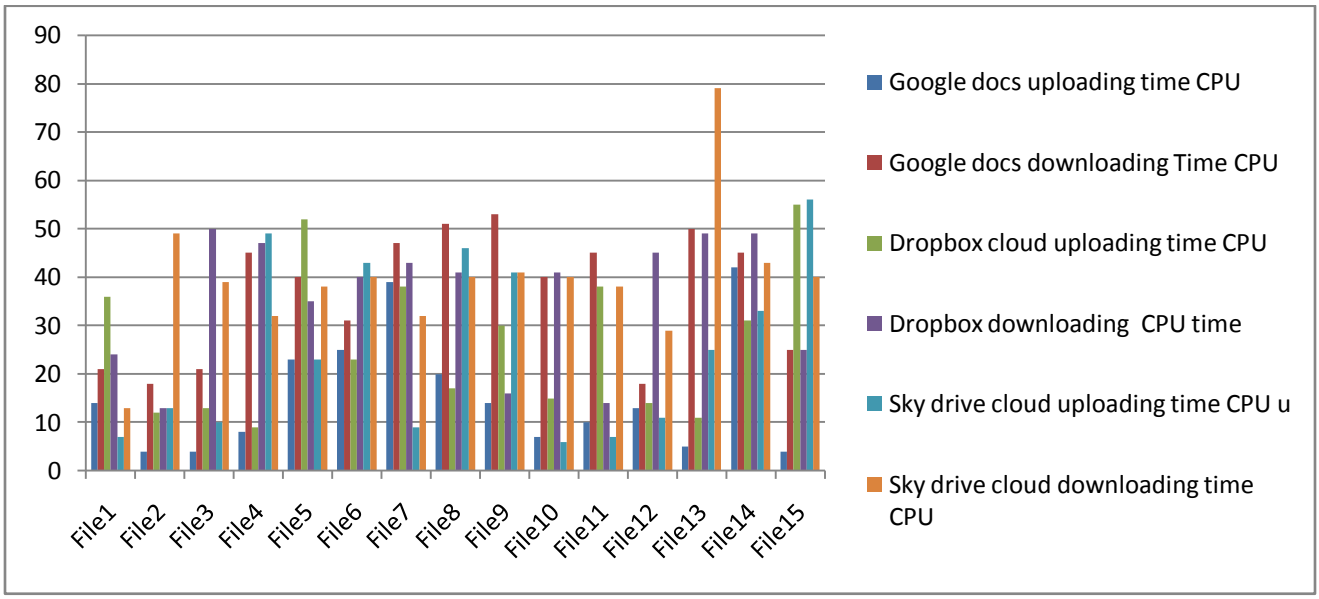
| **File name** | **Sky drive cloud uploading time CPU usage** | **skydrive uploading time memory usage** | **X1** | **Y1** | **Skydrive downloading Time CPU usage** | **skydrive downloading time memory usage** | **X2** | **Y2** |
|---|---|---|---|---|---|---|---|---|
| File1 | 7 | 1013.76 | 49 | 1027709 | 13 | 1024 | 169 | 1048576 |
| File2 | 13 | 1013.76 | 169 | 1027709 | 49 | 1013.76 | 2401 | 1027709 |
| File3 | 10 | 1003.52 | 100 | 1007052 | 39 | 1003.52 | 1521 | 1007052 |
| File4 | 49 | 1003.52 | 2401 | 1007052 | 32 | 1003.52 | 1024 | 1007052 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| File5 | 23 | 993.28 | 529 | 986605 | 38 | 993.28 | 1444 | 986605.2 |
| File6 | 43 | 1019.9 | 1849 | 1040204 | 40 | 995 | 1600 | 990025 |
| File7 | 9 | 1012.74 | 81 | 1025634 | 32 | 990 | 1024 | 980100 |
| File8 | 46 | 982 | 2116 | 964324 | 40 | 985 | 1600 | 970225 |
| File9 | 41 | 979 | 1681 | 958441 | 41 | 979 | 1681 | 958441 |
| File10 | 6 | 974 | 36 | 948676 | 40 | 974 | 1600 | 948676 |
| File11 | 7 | 967 | 49 | 935089 | 38 | 968 | 1444 | 937024 |
| File12 | 11 | 963 | 121 | 927369 | 29 | 963 | 841 | 927369 |
| File13 | 25 | 946 | 625 | 894916 | 79 | 969 | 6241 | 938961 |
| File14 | 33 | 940 | 1089 | 883600 | 43 | 953 | 1849 | 908209 |
| File15 | 56 | 934 | 3136 | 872356 | 40 | 944 | 1600 | 891136 |

The following chart gives graphical representation of CPU performance and memory usage over skydrive cloud.




From the above graphical representation it is clear that CPU performance while uploading and downloading fluctuates like dropbox cloud. Skydrive also shows fluctuation in system memory usage while uploading and downloading the files. The following chart gives the comparative analysis of CPU performance of different storage clouds as stated above.




## 4. CONCLUSIONS

From the above analysis it is clear that different storage clouds shows different behaviours in the performance of CPU and System memory usage. Google docs gives the stable output that is CPU time on Google docs is lesser than the downloading CPU time. And in regard to memory usage downloading file consumes more system memory as compare to uploading files.

## 5. ACKNOWLEDGEMENT

Authors are highly grateful to Dr. Gurdev Singh, Professor & Head, DCSE, Adesh Institute of Engineering & Technology, Faridkot for their precious guidance from time to time.

## 6. REFERENCES

[1] Gurdev Singh, Goruav Garg, Prince Jain, "Structure or Cloud Engineering" IJCA, Volume 33– No.8, November 2011.

[2] Manik Sharma, Chandni Sharma et. al. "Comparative Study of Static Metrics of Procedural and Object Oriented Programming Languages", IJOCT, February 2012.

[3] http://computer.howstuffworks.com/cloud-computing/cloud-storage1.htm (online)

[4] K. Zimmer ,voras, B. Mihaljevi, M. Orli,(2011) " Evaluating open source cloud computing soloutions" MIPRO 2011, May 23-27, 2011, Opatija, Croatia

[5] Manik Sharma, Gurdev Singh, Parneet Kaur, "A Comparative Study of Static object oriented Metrics", IJoAT, Vol3, No1. January 2012.

[6] Liladhar R. Raewatkar, Lanjewar, "Implementation of Cloud Computing", IJCA, Number 8 - Article 6, 2010

[7] Patrícia Takako Endo1, Glauco Estácio Gonçalves1,(2008) "A Survey on Open-source Cloud Computing Solutions" VIII Workshop em Clouds, Grids e Aplications.

[8] Michael Sabala, Robert L., Grossman, Yunhong Gu, (2008) "Compute and Storage Clouds Using Wide Area High Performance Networks National Center for Data Mining University of Illinois at Chicago January 31, 2008 PP-11.

[9] Nariman Mirzaei "Cloud Computing", IEEE International conference on cloud computing. 2008.