Data Sharing in the Cloud and Ensuring Accountability

Mailud Mohamed Mailud

M .Sc. Student. Department of Computer Science, Sam Higginbottom university of Agriculture, Technology& Sciences, Allahabad, U. P. – 211007, India.

Dr. W. Jeberson

Head of Department Computer Science, Sam Higginbottom university of Agriculture, Technology& Sciences, Allahabad, U. P. – 211007, India.

Salah Boubkar Salah

M.Tech.Student.Department of Eectronics and Communication Engineering, Sam Higginbottom university of Agriculture, Technology& Sciences, Allahabad, U. P. – 211007, India.

Abstract – In cloud computing environment resources are shared among various clients and it's important for system provider to allocate the necessary resources for the clients. As the sizes of IT infrastructure continue to grow, cloud computing is a new way of virtualization technologies that enable management of virtual machines over a plethora of physically connected systems. Cloud computing enables highly scalable services to be easily consumed over the Internet on an as-needed basis a major feature of the cloud services is that users' data are usually processed remotely in unknown machines that users do not own or operate. While enjoying the convenience brought by this new emerging technology, users' fears of losing control of their own data (particularly, financial and health data) can become a significant barrier to the wide adoption of cloud services. Here the multi-layered architecture is proposed to address accountability of the data while sharing in the multi user, heterogeneous and distributed computing environment. The multi-layered architecture is evaluated and shows that the accountability of the data is ensured which increases the trust between the end user and the service provider.

Index Terms – Cloud, Resources, Heterogeneous, Data Sharing.

1. INTRODUCTION

Background

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network .It has the potential to change the IT industry. It enables cloud customers to remotely store their data into the cloud so as to enjoy the on-demand high quality application and services from a shared pool of configurable computing resources. Cloud Computing is the result of evolution and adoption of all the existing technologies and paradigms. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. Clouds enable customers to remotely store and access their data by lowering cost of hardware ownership while providing robust and fast services. As Cloud Computing becomes prevalent, sensitive information are being increasingly centralized into the cloud. In this, data owners may share their outsourced data with a large number of users, who might want to only retrieve certain specific data files they are interested in during a given session. One of the most popular ways to do so is through keyword-based search. Such keyword search technique allows users to selectively retrieve files of interest and has been widely applied in plaintext search scenarios. In a cloud the service providers offer their resources as services to the general public. Public clouds offer several key benefits to service providers, including no initial capital investment on infrastructure and shifting of risks to infrastructure providers. However, public clouds lack fine-grained control over data, network and security settings, which hampers their effectiveness in many business scenarios.

Cloud Computing

Cloud computing , or the cloud ,as shown in figure (1.1) is a colloquial expression used to describe a variety of different types of computing concepts that involve a large number of computers connected through a real-time communication network (typically the Internet).cloud computing is a jargon term without a commonly accepted unequivocal scientific or technical definition. In science, cloud computing is a synonym for distributed computing over a network and means the ability to run a program on many connected computers at the same time. The phrase is also, more commonly used to refer to network-based services which appear to be provided by real server hardware, which in fact are served up by virtual hardware, simulated by software running on one or more real machines. Such virtual servers do not physically exist and can

therefore be moved around and scaled up (or down) on the fly without affecting the end user - arguably, rather like a cloud.


Figure (1-1) cloud system

The popularity of the term can be attributed to its use in marketing to sell hosted services in the sense of application service provisioning that run client server software on a remote location.

Cloud Share

The data storage process can be time-consuming and costly. This includes maintaining data servers, storage disks, firewalls, backup copies and disaster- recovery provisions. My Cloud Storage reduces these burdens, allowing you to store, retrieve, share, and analyze your data, day after day, without worrying about maintenance, scaling up or down or hardware and firmware u upgrades.
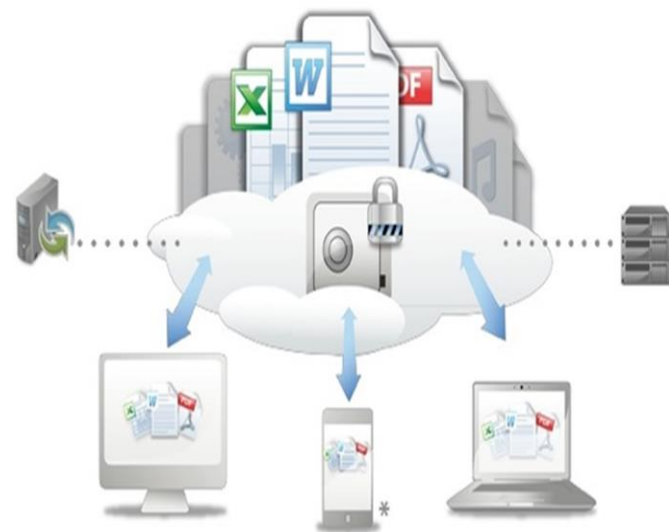
Cloud Share System

Proposed Work

Here the proposed design is a three layered architecture that will ensure the accountability and track the usage of owner's data. This architecture is developed to bring trust between end user and the owner regarding the usage of data. This architecture also enforces the proper handling of the owner's data according to the SLA (Service Level Agreement). Another advantage of this architecture is it can track the usage of data ,in future if any conflict arise then it can easily be trace down by the owner as well as the Cloud Service provider.

OBJECTIVES

1. To study the existing clouds.

2. To create a trusted cloud to store data & increasing the security level.

3. To analysis the developed cloud with existing cloud.

Research Methodology

Cloud computing is the delivery of computing services over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications.

The application perform the following operations-

Module Description – The following modules are introduced in to our application.

- User Module
- File Upload Module
- File Share Module
- Admin Module

User Module - Any person who wants to use the cloud storage service can register with the application. User registration is a simple process. After registration any one can login and can use the facilities like file upload, searching, sharing, changing profile images etc.

File Upload Module - Here the user can store their own data as well as can download the data whatever is uploaded by some other user. If you upload the data on the cloud then other users can search your information and can download after your confirmation.

File Share Module- Here any user can search your files from

the cloud storage then he/she will send a request on the owner for downloading the particular file. Then it is up to owner that he/she is allowing for that or not. If you not want to share your file then do not send the confirmation message, if you want for share you can send the confirmation message which contains a security key and by using that key any one can download your file.

Admin Module - The administrator can manage a registered user. The following tasks are performed by the administrator-

a. can block/unblock any registered user

b. can delete any registered user

c. can search any registered user

d. can send message to all the users of the cloud storage service

e. can see the message log

Concepts Used

Cloud Computing

When you store your photos online instead of on your home computer, or use webmail or a social networking site, you are using a "cloud computing" service. If you are an organization, and you want to use, for example, an online invoicing service instead of updating the in-house one you have been using for many years, that online invoicing service is a "cloud computing" service.

Cloud computing refers to the delivery of computing resources over the Internet. Instead of keeping data on your own hard drive or updating applications for your needs, you use a service over the Internet, at another location, to store your information or use its applications. Doing so may give rise to certain privacy implications.

Cloud Storage

Cloud Storage is a service where data is remotely maintained, managed, and backed up. The service is available to users over a network, which is usually the internet. It allows the user to store files online so that the user can access them from any location via the internet. The provider company makes them available to the user online by keeping the uploaded files on an external server. This gives companies using cloud storage services ease and convenience, but can potentially be costly. Users should also be aware that backing up their data is still required when using cloud storage services, because recovering data from cloud storage is much slower than local backup.

Advantages of Cloud Storage

Usability – All cloud storage services reviewed in this topic have desktop folders for Mac's and PC's. This allows users to drag and drop files between the cloud storage and their local storage.

Bandwidth – You can avoid emailing files to individuals and instead send a web link to recipients through your email.

Accessibility – Stored files can be accessed from anywhere via Internet connection.

Disaster Recovery – It is highly recommended that businesses have an emergency backup plan ready in the case of an emergency. Cloud storage can be used as a backup plan by businesses by providing a second copy of important files. These files are stored at a remote location and can be accessed through an internet connection.

Cost Savings – Businesses and organizations can often reduce annual operating costs by using cloud storage; cloud storage costs about 3 cents per gigabyte to store data internally. Users can see additional cost savings because it does not require internal power to store information remotely.

System Requirements Specification (SRS)

Introduction

This document aims at defining the overall software requirement for 'Ensuring Accountability for Data Sharing in the Cloud'. Efforts have been made to define the requirements exhaustively and accurately. The final product will be having only features/functionalities mentioned in this document.In case it is required to have some additional features, a formal change request will need to be raised and subsequently a new release of this document and/or product will be produced.

Purpose

This specification document describes the capabilities that will be provided by the software application 'Ensuring Accountability for Data Sharing in the Cloud'. It also states the various required constraints by which the system will put up with. The intended audiences for this document are the development team, testing team and end users of the product.

Scope

Ensuring Accountability for Data Sharing in the Cloud is a cloud storage service by which a user can share their files with each other as well as keep their important data on a cloud storage location.

Overview

The rest of this SRS document describes the various system requirements, interfaces, features and functionalities in detail.

Overall Description

Ensuring Accountability for Data Sharing in the Cloud provides a friendly environment and generally it is concerned to provide an efficient and effective way to give information online to its users.

Product Perspectives

The application will be web-based, self-contained and independent software product.
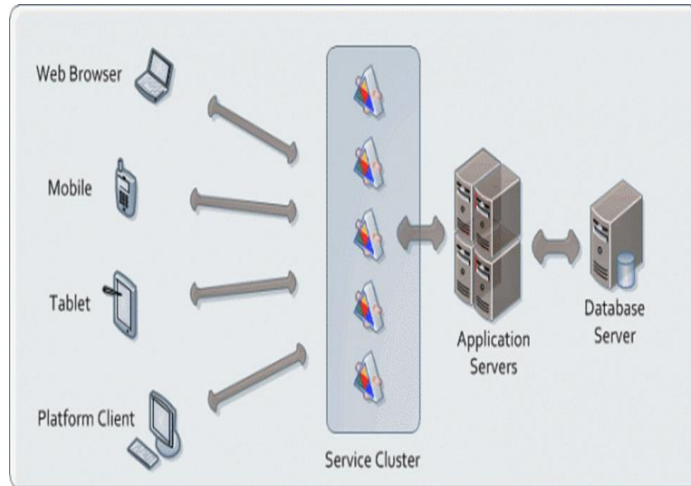
System Database Interaction

Methodology Adopted and Life Cycle of Project

A Software Life Cycle is the series of specialized stages that a software product undergoes during its lifetime. The first stage in the life cycle of any software product is usually the feasibility study stage. The subsequent stages are requirement analysis and specification, design, coding, testing and maintenance. Each of these stages is called a life cycle phase. A life cycle model represents all the activities required to make a software product transit through its life cycle phases.

It also captures the order in which these activities are to be undertaken. The strictest life cycle model used is the classical life cycle model. However, in any practical software development work, it is not possible to strictly follow the classical waterfall model from every phase to its preceding phase.

Methodology Used

- Structure System Analysis techniques had been adopted for the analysis of the algorithm and software development.

- Structure System Design techniques had been adopted for the design of the algorithm and software development.

- Prototyping Model had been used for the development of the GUI.

- Project Cost had been estimated as a function on the project, recourses acquired   for the project and COCOMO Model as the theoretical references.

- Test plan was created to form the strategy of testing this

includes the decision of testing techniques, decision of testing tools and decision of milestones when the testing will commences.

Analysis Methodology

- Structure System Analysis:

- Techniques are used for the analysis of algorithm and development.

- Back Ground Analysis

Concept of meaning and about the algorithm is learned through research Papers on internet and expert theory.

- Fact-Finding

Facts about the algorithm are collected from the books and internet.

- Fact Analysis

Facts are analyzed to through examples related to algorithm.

- DFD and Flow Chart

DFD and Flow Chart are used to model the algorithm and logic design.

Design Methodology

- Structure design techniques are used to design the software.
- An activity of all phases of project is the walk-through. It is the inter change of the ideas. In design walk-through the purpose is to recognize as many problems in the software as possible.
- Activity diagram are used then to define the behaviors of mechanism.
- Layering has been done to refine the architecture.
- Initially, two layers are introduced system layer and application layer.
- Database design is according to the implementation of algorithm.

Coding Methodology

- Even driven programming ASP.Net had been used for coding the modules and programs.

- Structured English and pseudo codes are used to closely refine the mechanism using the fatality of the define objects.

- Various steps had been used to facilitate incremental coding folloId by testing.

- The basic philosophy followed at this stage is code one line followed by rigoromeGoogletesting.

- Incremental compilation had been used to compile and

test on which work was in progress.

- Step wise refinement technique had been used to code the modules.

Testing Methodology

- Structured testing tools had been used for testing the programs.

- Small programs are developed to test the individual modules.

- Each module is tested by the development of appropriate functionality.

- Some other users assign to test the software functionality.

- The same users responsible to test the reliability of software by using variomeinput techniques, on variomeparts and functionality.

- The users had the Alpha Testing.

- Software distributed to users.

Modularization Details

System Analysis

This is one of the phases of SDLC (Software Development Life Cycle). It is a detailed appraisal of the existing system and includes how the system works and what it does; it also includes finding out more details what the system problem are and what are management techniques, which helps me in designing a new system and improving the existing system.

Elements of System Analysis

There are 4 basic elements in the system analysis. These are the following:-

a) Outputs:-

First I must determine the objective or goal what I intend to achieve, what is the purpose of my work, in other words, what is the main aim behind the system.

Defining aim is very vital in system work. If I do not know where I want to go, I will never know how to reach there. I shall be wasting unnecessarily time and energy in the purpose. Once I know my aim, I can try to achieve it in best possible way. The user has to define the objective in their terms and needs. These become outputs, which the analyst has to keep in mind.

b) Inputs:-

Once I know the output, I can easily determine what the inputs should be. Sometimes, it may happen that the required information may not be readily available in proper form. This may be because the existing forms are not properly designed. Sometimes, it may not be possible to get required information

without the help of top-level management. If the information is vital to the system, I should make all possible efforts to make it available.

Sometimes, it might be too costly to get the required information.

Essential elements of inputs are:-

1). Accuracy:-If the data is not accurate, the output will be wrong.

2). Timeliness:-If the data not obtained in time, the entire system falls into arrears.

3). Proper Format:- The inputs must be available in the proper format.

4). Economy:- The data must be produced at the least cost.

c). Files:

As the word implies files are used to store data. Most of the input necessary for the system may be historical data; or it may be generated from within the system; these are stored in files either in isolated form or in large volumes.

d). Processes:

This involves the programs in which data is processed through the computer. The processing involves the set of logical steps. These steps are required to be instructed to the computer and this is done by series of instructions called "Programs".

System Design

The system design is not a step adherence of clear procedures and guidelines. Though certain clear procedures and guidelines have emerged in recent days, but still much of design work depends on knowledge and experience of the designers.

The scope of the system design is guided by the framework for the new system developed during analysis. More clear defined logical method for developing system that meets user requirements has led to new techniques and methodology that fundamentally attempts to do the following:-

- Improve productivity of analyst and programmer.

- Improve documentation which means subsequent enhancements.

- Cut down drastically on cost over runs and delays.

- Improve communications among the users, analyst, designers, and programmers.

- Standardize the approach to analysis and design.

- Simplify design by segmentation.

2. RESULTS AND DISCUSSION

Execution process:

i. Before using the cloud service the user must register.

ii. After registration the user can login.

iii. Now the user can upload their file to the cloud storage with a keyword (keyword specifies the type and category of the information which belongs to the uploaded file).

iiii. The user can also search the information which is uploaded by some other user.

v. If the required data is available so the logged in user can request for download.

vi. Now the next user can confirm the download request, if he/she is agreed for sharing the data.

vii. If download request is confirm then a confirmation message will send to the request user.

viii. In the confirmation message there is a download link, by clicking this link and providing the security key the user can download the particular file.

ix. User can also change their profile photo if he/she wants.

x. We also work on Admin part, the administrator can delete user, block user or delete the user.

xi. The Admin can also broadcast the message to the entire registered user of cloud storage service.

Efficiency of Cloud Storage:

Approximate all the mail services provides the facility for attaching a file along with the mail, so that we can send their files to the target user. Now the question is – what is the use of this storage service? So, any mail service provides a limited size file for attachment (normally 20 MB maximum). But here file size is no bar; you can attach any file without thinking about the size of the file.

Discussion

i. The project is executed with high security, without login the user not access any kind of service.

ii. The logged in user can not download a file directly. After confirmation of the owner of the file the logged in user can download with a security key.

iii. The system has been developed by using ASP.NET. It provides a lot of facility for designing a good user interface as well as server side programming.

iiii. The data has implemented by using SQL-SERVER.

v. Here we also use AJAX for designing the interactive user interface.

3. CONCLUSION

In this dissertation, we implement a secured cloud data share system, by which any registered user can share their data with other registered user's of the cloud storage service. I try to maintain a security system, when any user wants to download the uploaded data the he/she should first take the permission of the owner of the data. Only when the owner of data is agreed then the user can download the data with the help of the security key send by the owner of the data..

Here I also try to maintain a admin part which is the management console of the cloud storage. The administrator has authority to manage the user and block, unblock etc.

Future Enhancements

I have developed this application for the web servers and this is most suitable for personal computers. In future it could be also designed for android based mobile phones and tablets. We can also do more work on the security features of the application. May be send the security code on the mobile phone of the registered users. This concept will increase the security of the system. We can also use some advance concepts like MVC and JQuery for enhancing the performance of the application.

REFERENCES

[1] MdMasoomRabbani, Enhancing Accountability for Distributed Data Sharing in the Cloud, International Journal of Engineering and Technology (IJET), Vol 5 No 3 Jun-Jul 2013, ISSN : 0975-4024

[2] Anup Mathew, Survey Paper on Security & Privacy Issues in Cloud Storage Systems, EECE 571B, TERM SURVEY PAPER, APRIL 2012

[3] Sudha S and Brindha K, Data Synchronization Using Cloud Storage, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 11, November 2012, ISSN: 2277 128X

[4] TalasilaSasidhar, Pavan Kumar Illa, SubrahmanyamKodukula, A Generalized Cloud Storage Architecture with Backup Technology for any Cloud Storage Providers,INTERNATIONAL JOURNAL OF COMPUTER APPLICATION,ISSUE2, VOLUME 2 (APRIL 2012) ISSN: 2250-1797

[5] Wenying Zeng, Research on Cloud Storage Architecture and Key Technologies,ICIS 2009, November 24-26, 2009 Seoul, Korea Copyright © 2009 ACM 978-1-60558-710-3/09/11... $10.00"

[6] Victor Chang, Robert John Walters, Gary Wills, Cloud Storage in a private cloud deployment: Lessons for Data Intensive research

[7] Ronny Seiger, SecCSIE: A Secure Cloud Storage Integrator for Enterprises

[8] Eric A. Fischer, April 23, 2013, Overview and Issues for Implementation of the Federal Cloud Computing Initiative: Implications for Federal Information Technology Reform Management

[9] Mladen A. Vouk, Journal of Computing and Information Technology - CIT 16, 2008, 4, 235–246, Cloud Computing – Issues, Research and Implementations

[10] IdilioDrago, Inside Dropbox: Understanding Personal Cloud Storage Services, IMC'12, November 14–16, 2012, Boston, Massachusetts, USA. Copyright 2012 ACM 978-1-4503-XXXX-X/12/11 ...$15.00.