File Sharing System

	RapidShare	4Shared	Mediafire	F-Secure
SMS Verification	No	No	No	Yes
Key Generation	No	No	No	Yes
Traffic/Bandwidth	Limited	Limited	Limited	Unlimited
Link Generation	Yes	Yes	Yes	Yes

[1]Ajeet Kumar Singh, [2]Deepak Singh, [3]Sagar Goswami, [4]Prof. Bandana Jha

Department Of Information Technology

Abstract: File sharing is the public or private sharing of computer data or space in a network with various levels of access privilege. While files can easily be shared outside a network (for example, simply by handing or mailing someone your file on a diskette), the term file sharing almost always means sharing files in a network, even if in a small local area network. File sharing allows a number of people to use the same file or file by some combination of being able to read or view it, write to or modify it, copy it, or print it.

Typically, a file sharing system has one or more administrators. Users may all have the same or may have different levels of access privilege. File sharing can also mean having an allocated amount of personal file storage in a common file system

Index Terms: Secure data transfer component, Key generation

I. INTRODUCTION

The use of the terms uploading and downloading often imply that the data sent or received is to be stored permanently, or at least stored more than temporarily. In contrast, the term downloading is distinguished from the related concept of streaming, which indicates the receiving of data that is used near immediately as it is received, while the transmission is still in progress and which may not be stored long-term, whereas in a process described using the term downloading, this would imply that the data is only usable when it has been received in its entirety. Increasingly, websites that offer streaming media or media displayed inbrowser, such as YouTube, and which place restrictions on the ability of users to save these materials to their computers after they have been received, say that downloading is not permitted. In this context, download implies specifically "receive and save" instead of simply "receive".[1] However, it is also important to note that downloading is not the same as "transferring" (i.e., sending/receiving data between two storage devices would be a transferal of data, but receiving data from the Internet would be considered a download.

Uploading" and "downloading" are terms used to refer to types of electronic data transfers. The difference between them is the direction in which the files are being transferred. Files are considered to be uploaded when they are transferred from a computer or other electronic device to a central server, and downloading is when the files are transferred from a server to a smaller peripheral unit, such as a computer, smart phone or other device. [2] These two different types of transfers are often done via the Internet, such as when a file is downloaded from a website. The transfer of data from one system or device to a similar system or device, such as from a desktop, computer to usually is not considered uploading or downloading.

II. RELATED WORK

Currently the system used for storing and again retrieving the data is very costly and used at a large scale. No matters, the systems provide many operations along with data storage but they are paid system. So because of this major disadvantage of the currently available systems, it is used by the premium users only i.e. paid users.

In the current scenario for such systems it is necessary that you must provide your credit card number, as it is paid service. Some systems also provide the facility of free data storage and retrieval but it is for limited number of days only. Currently system does not provide any authentication criteria for retrieving the data. There are many such problems with the existing system. The main drawbacks of existing system are:

	RapidShare	4Shared	Mediafire	F-Secure
SMS Verification	No	No	No	Yes
Key Generation	No	No	No	Yes
Traffic/Bandwidth	Limited	Limited	Limited	Unlimited
Link Generation	Yes	Yes	Yes	Yes

-        Expensive: Only paid users are allowed to send the data also user has to submit the Credit card number to pay for respective services.

- Less secured: Current systems do not provide any full-proof mechanism for authentication of the user.
- Unreliable: As no full-proof mechanism for user authentication is supported by currently used systems, therefore it is not reliable to make use of such services or software's.

III. PROJECT SCOPE

On domestic level where students are concerned it could be used as software where in it is possible to transfer power point presentation, assignments, programs, video clips could be sent within few of minutes. On industrial level where data privacy/security is a major concern, our software provides efficient security and adding it up with unlimited data transfer.

3.1 User Classes and Characteristics:

This software applies to any user of any computer system. There is no limitation on who would be able to or allowed to use F-Secure, as F-Secure is an application that could be used by user of a computer system who would like to manage as well as Upload and Download data i.e. files on network. On domestic level this application is useful in Schools, Colleges, Universities where Students can make use of this software to get required data (assignments, practical's etc). On industrial level many companies can use this software to transfer their highly sensitive data in other companies. *[3]*

- Administrator: They are one of the core users and are able to authenticate the client who is going to upload and download data.
- Characteristics: Administrator must have knowledge of working of file sharing systems and he/she should be able to solve the problems that come across during the file transfer process.
- End User: The end users are expected to be Internet literate because they are the people who are going to get the key and verification message.
- Characteristics: No technical experience is required basic knowledge of handling system is sufficient.

3.1.1 Operating Environment

As far as operating systems are concerned F-Secure is compatible with:

-All 64-bit versions of Microsoft Windows

-All 32-bit Microsoft Windows (95/98/NT/2000/XP/Vista)

-All POSIX (all versions of Linux/all versions of BSD/all UNIX-like operating systems)

Moreover, F-Secure is produce with the help of open source both front end and back end that is compatible with almost any operating system. It is obvious that F-Secure is independent from the operating system of the computer system on which it runs. This software is web-based and can be viewed by any web browser, i.e. Mozilla, Internet Explorer, Netscape Navigator, and Opera.

3.1.2 Design and Implementation Constraint

Main server should always be on. Every client system should have web browsers like Google Chrome, Mozilla Firefox, Internet Explorer, Netscape Navigator, and Opera or web browsers which are compatible to above mentioned browsers. Every system should also have flash player installed in it.

Anyone who wishes to work on further development of FSecure should have the knowledge of following programming languages: • HTML (Front) •Java Script (Front) •Nodejs, mongo dB, ejs (Back).

3.1.3 Assumptions and Dependencies

Every user should provide the mobile number compulsorily, because the authentication code will be sent to the same mobile number. Every new visitor should fill registration form and set the user id and password. While downloading the data, the user should enter correct file password, authentication key and Captcha code. We assume that extra documentation beyond this report would not be necessary in order for the user to utilize this software.

3.3 External interfaces Requirements

3.3.1 User Interfaces

The first interface is the log-in screen. This is where the user has a specific Username and Password so that they can gain access to the software and data. User can also choose that whether he/she wants to upload or download the data. One of F-Secure's advantages is its user interfaces, as they are extremely functional, easy to use and can be handled even by users with very little knowledge and experience on using computer systems.

3.3.2 Hardware Interfaces

Though not necessarily interfacing with the hardware, the system must make use with an internet connection. For FSecure's installation and use, requirements are almost zero. It can run on any computer system, regardless of the type of operating system used. The only thing F-Secure require, as far as hardware is concerned, is a compatible CPU.

3.3.3 Software Interfaces

Along with the internet connection, the system makes indirect use of an internet browser. Outside of the HTML code and PHP, the code doesn't tell any software, including the browser, what to do.

-HTML: It is used to develop various pages and forms.

-Nodejs: It is used to carry data.

-JavaScript: It is used to perform logical calculations and looping activities.

-CSS: It is used to place the components on pages and forms.

Moreover, F-Secure is produce with the help of open source both front end and back end that is compatible with almost any operating system. It is obvious that F-Secure is independent from the operating system of the computer system on which it runs. This software is web-based and can be viewed by any web browser, i.e. Mozilla, Internet Explorer, Netscape Navigator, and Opera.

3.3.4 Communication Interfaces

The system uses an internet connection to connect to the database. The code itself though, does not specifically direct the network controllers to do any work. Connection to the Internet would be considered useful for users, as, through it, they would be able to access the software and information about F-Secure.

-Mongo dB: Multiple systems communicate with each other using common database stored on main server and also it is an open source and feely available.

-Email-to-email for data transmission.

3.4 Non-Functional Requirements:

3.4.1 Performance

For system to perform precisely or affectively bandwidth is the major requirement. Even if minimum requirement is around 256KBPS, system to run smoothly, must be provided bandwidth of around 1MBPS.

3.4.2 Safety

To keep the uploading/downloading data safe, there should not be any voltage fluctuations of computer system. This application does not run in cases of wrong pass word/key insertion or wrong settings. In case of error it provides an alert message to users with appropriate help messages.

3.4.3 Security

System generated key i.e. private key must be kept hidden from any other person except the One on receiving end. Access to the database should be restricted to people who are not valid/authenticated users. Passwords and ID's should be regulated to be at least a certain length and also contain alphanumeric characters in both the password and key generation.

3.5 Other Requirements 3.5.1 DB requirements:-

MongoDB is a document-oriented NoSQL database used for high volume data storage. Instead of using tables and rows as in the traditional relational databases, MongoDB makes use of collections and documents. Documents consist of key-value pairs which are the basic unit of data in MongoDB. Collections contain sets of documents and function which is the equivalent of relational database tables. MongoDB is a database which came into light around the mid-2000s.

Features:

1. Each database contains collections which in turn contains documents. Each document can be different with a varying number of fields. The size and content of each document can be different from each other.
2. The document structure is more in line with how developers construct their classes and objects in their respective programming languages. Developers will often say that their classes are not rows and columns but have a clear structure with key-value pairs.
3. The rows (or documents as called in MongoDB) doesn't need to have a schema defined beforehand. Instead, the fields can be created on the fly.
4. The data model available within MongoDB allows you to represent hierarchical relationships, to store arrays, and other more complex structures more easily.
5. Scalability – The MongoDB environments are very scalable. Companies across the world have defined clusters with some of them running 100+ nodes with around millions of documents within the database Key Components of MongoDB Architecture Below are a few of the common terms used in MongoDB

1. **_id** – This is a field required in every MongoDB document. The _id field represents a unique value in the MongoDB document. The _id field is like the document's primary key. If you create a new document without an _id field, MongoDB will automatically create the field. So for example, if we see the example of the above customer table, Mongo DB will add a 24 digit unique identifier to each document in the collection.
2. **Collection** – This is a grouping of MongoDB documents. A collection is the equivalent of a table which is created in any other RDMS such as Oracle or MS SQL. A collection exists within a single database. As seen from the introduction collections don't enforce any sort of structure.
3. **Cursor** – This is a pointer to the result set of a query. Clients can iterate through a cursor to retrieve results.
4. **Database** – This is a container for collections like in RDMS wherein it is a container for tables. Each database gets its own set of files on the file system. A MongoDB server can store multiple databases.
5. **Document** – A record in a MongoDB collection is basically called a document. The document, in turn, will consist of field name and values.
6. **Field** – A name-value pair in a document. A document has zero or more fields. Fields are analogous to columns in relational databases.The following diagram shows an example of Fields with Key value pairs. So in the example below CustomerID and 11 is one of the key value pair's defined in the document.
7. **JSON** – This is known as JavaScript Object Notation. This is a human-readable, plain text format for expressing structured data. JSON is currently supported in many programming languages.
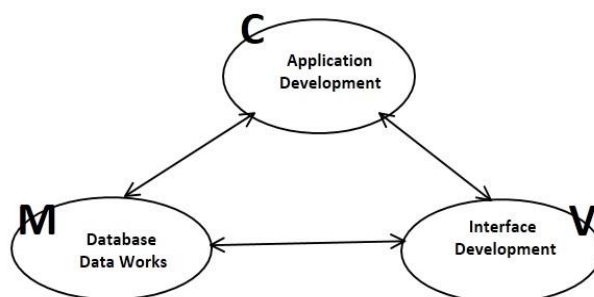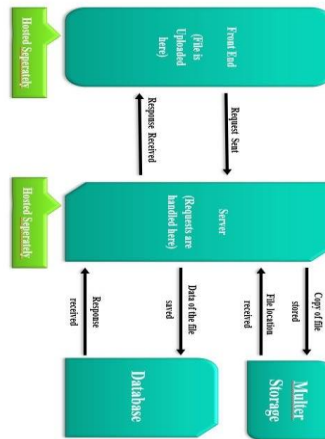
IV. SYSTEM ARCHITECTURE

4.1 MVC

Fig. 4.1: System flow

As you see above if you split your project three part like Application Development, Database and Data processing and Interface development then of course you project will be more understandable and developers can work more efficiently.

 Model: It handles data processing and database works part. Model processes events sent by controller. After processing these events then it sends processed data to controller (thus, controller may reprocess it) or directly to view side.

-          View: View prepares an interface to show to the user. Controller or model tells view what to show to the user. Also view handles requests from user and informs controller.

Controller: Let's say controller is like brain of the system. That is right. Because it processes every request, prepares other parts of the system like model and view. Then the system determines what to do by controller's commands. Even if you are junior developer you probably faced with a problem called "Complexity of Project". There MVC comes.

-          MVC helps you to decrease complexity of project. For instance, there is a team with 5 people. If they are working on same project after sometime project reaches a point which no one can understand what project is going to be. MVC splits project to (by default) three different parts. Like Model, View, Controller.[4]

4.2 **CodeIgniter**

   CodeIgniter is an open source rapid development web application framework, for use in building dynamic web sites with PHP. "Its goal is to enable [developers] to develop projects much faster than...writing code from scratch, by providing a rich set of libraries for commonly needed tasks, as well as a simple interface and logical structure to access these libraries."

   CodeIgniter is loosely based on the popular Model-ViewController development pattern. While view and controller classes are a necessary part of development under CodeIgniter, models are optional.

   Features:
- Extremely Light Weight
- Full Featured database classes with support for several   platforms.
- Form and Data Validation
- Session Management
- Email Sending Class. Supports Attachments, HTML/Text   email, multiple protocols and more - File Uploading Class
  Search-engine Friendly URLs

 V. TECHNICAL SPECIFICATION

 The technologies which are used to implement the system are:

 1) HTML is used for webpage creation.

 2) PHP is used as server side language as it is free and open source.

 3) Java Script and JQuery for webpage validation and dynamic generation.

 4) Eclipse PHP is used as a Rapid Application Development
     Tool (RAD) or as an Integrated Development Environment (IDE) for coding the software.

 5) MySQL is used as database management system because it is free and open source and also fast.

 Back-End:

 An http server is created in the backend, that helps to monitor and react based on the request from the frontend.

 The API contains different routes for handling of different tasks like uploading, downloading and processing/displaying of files. These routes serve different requests like POST, GET, etc. which based on the HTTP middleware. The routes help to manage the task individually.

 The API is connected with the cloud mongo cluster that contains the collections of all the files uploaded till now, and accordingly maintains a schema for the same.

 After each 24 hours, the server sends a command to the cloud to wipe the data for a fresh start for the next day, this in turn maintains the privacy of the content shared, as we delete it at the end of the day.

 5.1 Advantage

-          Security: Our site provides highest security for file hosting by exact authentication of downloader via SMS verification and key matching.

-          Unlimited Bandwidth and traffic: Our site gives unlimited data storage and bandwidth it also handles large traffic of data.

-          CAPTCHA Verification: For every download we used CAPTCHA verification so that various attack on site can be prevented.

5.2 Disadvantage
- SMS verification would not work if mobile phone is out of range.
- If downloader lost his mobile phone then it is almost impossible for him to get the file without phone.

REFERENCES

1. Architecture of the Secure File System, James P. Hughes, Storage Technology Corporation, jim@network.com, Christopher J. Feist, Storage Technology Corporation chris.feist@network.com.
2. [2]Enhancement of anticipative recursively adjusting mechanism for redundant parallel file transfer in data grids.
3. Peter Gutmann, University of Auckland, New Zealand. The secure filesystem (sfs) for dos/windows. http://www.cs.auckland.ac.nz/ pgut001/sfs/index.html, September 1996.
4. Alex Tormasov. The tordisk project. http://www.cs.auckland.ac.nz/ pgut001/sfs/index.html, November 1997.
5. Matt Blaze. A cryptographic file system for UNIX. In *First ACM Conference on Communications and Computing Security*, pages 33–43, Fairfax, VA, November 1993.