**DOCSTASH**

**(Personal Storage Solution)**

Submitted in partial fulfillment of the requirements

for the degree of

**Bachelor of Engineering**

**in**

**Information Technology**

by

**Rakshan S. Shetty 12IT1017**

**Rohit N. Ambre 13IT3002**

**Nakul I. Chauhan 13IT1049**

**Shakeel S. Shaikh 12IT1096**

Supervisor

**Prof. Nilima Dongre**



Department of Information Technology

Dr. D. Y. Patil Group’s

**Ramrao Adik Institute of Technology**

Nerul, Navi Mumbai 400706**.**

(Affiliated to University of Mumbai**)**

( 2016)



Ramrao Adik Institute of Technology

(Affiliated to the University of Mumbai)

Dr. D. Y. Patil Vidyanagar, Sector 7, Nerul, Navi Mumbai 400706.

CERTIFICATE

*This is to certify that, the dissertation titled*

“ DOCSTASH

(Personal Cloud Storage) ”

*is a bonafide work done by*

Rakshan S. Shetty

Rohit N. Ambre

Nakul I. Chauhan

Shakeel S. Shaikh

*And is submitted in the partial fulfillment of the requirement for the degree of*

Bachelor of Engineering

in

Information Technology

to the

University of Mumbai



Supervisor

Prof. Nilima Dongre

Project Co-ordinator Head of Department Principal

(Prof. Reshma Gulwani) (Prof. Ashish Jadhav) (Dr.Ramesh Vasappanavara)

Project Report Approval for B.E.

This is to certify that the project entitled *"DOCSTASH (Personal Cloud Storage)”* is a bonafide work done by *Rakshan S. Shetty, Rohit N. Ambre, Nakul I. Chauhan and Shakeel S. Shaikh* under thesupervision of *Prof. Nilima Dongre***.** This project has been approved for the award of *Bachelor’s Degree in Information Technology, University of Mumbai.*

Examiners:

1...............................

2...............................

Supervisors:

1...............................

2.............................

Principal :

..............................

Date :

Place :

Declaration

I declare that this written submission represents my ideas in my own words and where others ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Rakshan S. Shetty 12IT1017 \_\_\_\_\_\_\_\_

Rohit N. Ambre 13IT3002 \_\_\_\_\_\_\_\_

Nakul I. Chauhan 13IT1049 \_\_\_\_\_\_\_\_

Shakeel S. Shaikh 12IT1096 \_\_\_\_\_\_\_\_

Date:

**Abstract**

Cloud computing provides an enormous amount of virtual storage to the users. Cloud storage mainly helps to small and medium scale industries to reduce their investments and maintenance of storage servers. Cloud storage is efficient for data storage. Users' data are sent to the cloud is to be stored in the public cloud environment. Security of cloud storage is ensured through confidentiality parameter. To ensure the confidentiality, the most common used technique is encryption. To have efficient cloud storage confidentiality. Based on the type of data, encryption and can be applied. Confidentiality can be achieved with a combination of encryption. The user data is encrypted using AES and divided into multiple blocks and stored on different cloud servers.

“Docstash” Application handles and manages the day to day data storage activities. It is a cloud storage solution provided to user where data can be stored in the cloud environment. Stored data is encrypted and compressed where users can access from anywhere. User can modify his data in Docstash without having the actual data in his physical device.

This project aims at providing a generic platform for cloud storage solution. DOCSTASH is a platform using latest technologies to provide a cloud environment that managed data integrity and security.

i

List of Figures

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Name** | **Page No.** |
| Fig 1.1 | Concept of Cloud Storage | 1 |
| Fig 4.1 | System Architecture of DOCSTASH | 9 |
| Fig 5.1 | Activity diagram of DOCSTASH | 12 |
| Fig 5.2 | Level 0 Data Flow Diagram of User Information | 13 |
| Fig 5.3 | Level 0 Data Flow Diagram of File storage | 13 |
| Fig 5.4 | Level 1 Data Flow Diagram of File Storage | 14 |
| Fig 5.7 | User interface | 15 |
| Fig 5.7 | Development Tool | 16 |

ii

ii

Contents

Abstracti

List of Figuresii

1. Introduction1

1.1 Concept of Cloud Storage1

1.2 Problem Definition2

1.3 Scope of project2

1.4 Relevance and Motivation of project3

1.5 Organization of Report3

2. Review of Literature5

3. Planning and Formulation6

4. Methodology8

4.1 Proposed Methodology8

4.2 System Requirements9

5. Design of System11

5.1 System Design11

5.2 Data Flow Diagrams13

6. Conclusion and Future Scope17

Literature Cited18

Acknowledgement19

1. **Introduction**
   1. **Concept of Cloud Storage**

Cloud Storage is a model of data storage in which the digital data is stored in logical pools, the physical storage spans multiple servers (and often locations), and the physical environment is typically owned and managed by a hosting company. These cloud storage providers are responsible for keeping the data available and accessible, and the physical environment protected and running. Cloud storage services may be accessed through a co-located cloud computer service, a [web service](https://en.wikipedia.org/wiki/Web_service) [application programming interface](https://en.wikipedia.org/wiki/Application_programming_interface) (API) or by applications that utilize the API, such as cloud desktop storage, a [cloud storage gateway](https://en.wikipedia.org/wiki/Cloud_storage_gateway) or [Web](https://en.wikipedia.org/wiki/World_Wide_Web)-based [content management systems](https://en.wikipedia.org/wiki/Content_management_systems).

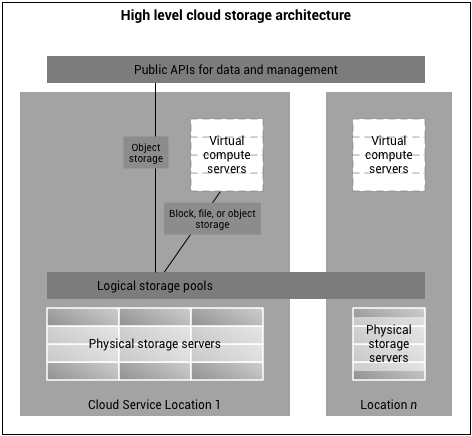


Fig 1.1 Concept of Cloud Storage

1

A Cloud storage has the following three main characteristics:

• View and edit files from the cloud

• Security and file encryption

• Flexible storage capacity at an affordable price

• Reliable tech support

Our project, ”**DOCSTASH – Personal cloud storage**” is based on above features which explores Cloud Storage in depth.

* 1. **Problem Definition**

To keep up with the ever-expanding digital data of the user, storing and managing those files are difficult. The main problem that arises when dealing with storage is data security and data integrity. This files should get corrupt while transferring to the cloud. The user also needs to have access to this files immediately and convenience of file access is also important. Also, unauthorized users should not have access to user files. Moreover, the user should have a way to share this files with others easily.

**DOCSTASH** attempts to bring cloud security to user’s data where they can store files reliably, proposing a new way of storage where data integrity and data confidentiality will be maintained.

It is based on zero knowledge concept, where user information and data is encrypted and stored in database

* 1. **Scope of Project**

This project aims at providing a generic platform to enhance storage management and provides integrity by encryption. Users can manage and modify their data by very user friendly interface.

Each user will have storage capacity of 10 GB where they can save any file format. While the music and videos can be streamed directly via browser. All the user data will be encrypted automatically and data backup will be taken periodically.

2

Our project is designed to achieve the following targets:

**• User authentication**: It provides better user authentication through “JWT authentication”.

**• Zero knowledge**: It uses zero knowledge concept to store data on cloud environment. efficiently.

• **Direct download**: It provides direct download of online content to user’s account.

**• Torrent download**: Users can search for torrent files or can provide magnet links and it will download the files in cloud storage.

• **Storage expansion**: Users can expand their storage space by changing their plan to Pro version.

* 1. **Relevance and Motivation of Project**

As more numbers users are getting access to world of internet need of data storage is required. Users need access to more amount of data daily than what they can store in their physical devices, carrying big amount storage devices is not convenient for commuters. Storing such data on cloud environment can provide easier access to any stored data.

Users want to have security and integrity to their data as data security is usual concern of data on cloud environment. Users want to make sure that their data must be kept in secure cloud environment. Traditional cloud storage providers do not allow user to expand memory which is allocated to them. However, some of the providers are trying to implement a platform so that they can fulfill user’s requirement and resolve the problem for the same.

Therefore, we need to develop a platform that provides enormous amount of data to be kept securely with providing additional cloud storage.

* 1. **Organization of the Report**

**Chapter 1:**

The primary objective of this chapter is to introduce the reader about problem definition, scope, and motivation behind this project

3

**Chapter 2:**

This chapter represents a critical appraisal of previous work related to the project work

**Chapter 3:**

This chapter deals with planning and formulation of the project which explains the timeframe and modules of the project in detail.

**Chapter 4:**

This chapter includes the final outcome of the project and the different methods employed to achieve the same. It also includes the system requirements needed for the project

**Chapter 5:**

This chapter details the system design including working model and its implementation. This also includes Data Flow Diagrams (DFD) and activity diagram of the project.

**Chapter 6:**

This chapter defines conclusion, future scope and references of the project

4

1. **Review of Literature**

Prof Feige, U., Fiat, A. & Shamir in their paper introduced to the concept of zero knowledge which was further implemented by Springer Inc. It gives maximum protection to users data by client-side encryption technology to secure users file in their own devices with some high-grade encryption and the files cannot be decrypted in a cloud.

Digital Rights Management was introduced by Tresorit Inc in their paper where they explain the process of file sharing in collaboration process. The DRM defines which user have right to modify the file or just read the file so that users with permission can only change the content of files while sharing.

Using different types of storage system as per need is described by Prof Arun Teneja. File and block storage might provide better performance, but granular metadata and near-infinite scalability make object storage equally beneficial. Use cases for each storage system is described in his paper.s

“Google Drive” is a cloud storage solution by Google. In Drive users get their storage when they sign in for a new account.

Drive also has some unique features that integrate different third party software in it. This software helps us in editing and managing our files and folders online. It help in editing those file in collaboration mode so multiple users can edit simultaneously.

Dropbox is another such cloud storage provider with unique features where users can download the native desktop application and the files and folder in specific location will automatically be synchronized. They also provide enterprise support for storage of company’s large amount of data. Dropbox users as of March 2016. As of that month, the cloud-storage service had 500 million registered users, up from 300 million in May 2014.

5

1. **Planning and Formulation**

Developing an SaaS platform is a tedious task. It takes long time and big efforts to develop a cloud storage based web application.

The aspects included while developing cloud storage platform are:

* Creating Cloud environment
* Implementation of latest technologies
* Strong User Authentication
* Implementing security methods
* Errorless cloud storage platform

To manage all these critical components of project planning is must.

Firstly, we started with searching, to know what other storage providers are offering and the techniques used by them. Understanding their flow of working and different services offered to decide required services in our project.

**Time duration: 2 week**

Then we started to make static web pages for docstash by analyzing interfaces of all other storage providers. To make web pages we started by sketching designs of web pages and then implementing it into real web pages, Also understanding way to establish back-end support.

**Time duration: 45 days**

The next step is providing strong User authentication methods. This is crucial step in cloud application or any software to ensure data integrity and confidentiality. To provide more secure access we worked on Third party user authentication. Implementing latest front and back end techniques to make application fast and dynamic.

6

**Time duration: 45 days**

Implementing security features to provide more reliable platform. Providing encryption to files stored on cloud to secure it from any threat. Testing of application to find any errors or bug to deploy totally bug free platform for users. All the steps mentioned above are merged in this final step. The most important step in this project is creating our own cloud environment, adding the data to the database, applying security methods in this critical step.

**Time duration: 45-60 days**

7

1. **Methodology**
   1. **Proposed Methodology**

* Docstash is  SaaS based cloud storage platform to solve the lack of physical storage space in devices.
* It will ensure data integrity and security of user files
* It allow user to view edit and download their files from any devices.
* User can login with 3rd party service like Google to extend their storage.
* All files will be encrypted with industry grade AES-256 Algorithm and stored and served with compression enabled.
* Users can use BitTorrent protocol to download files directly to their storage
* Users can have access to social media platform like Youtube, Twitter even if there ISP has restricted access to such websites.
* Docstash will provide assistance to a user with user-friendly bot always present to help.
* User and can interact with another user through a chat interface and share. link to friends which is shortened via link shortner built  into docstash.
* Docstash will be built upon new and cutting edge software stack that will enable it to scale easily.
* Every component will run on different containers which will help to scale each component independently of other.
* With the Non-Blocking Restful API users will get their responses from server very fast.
* All updates to the user files will be provided to a user in Real Time.
* Multi-Tenant user supports multiple devices at same time.

8

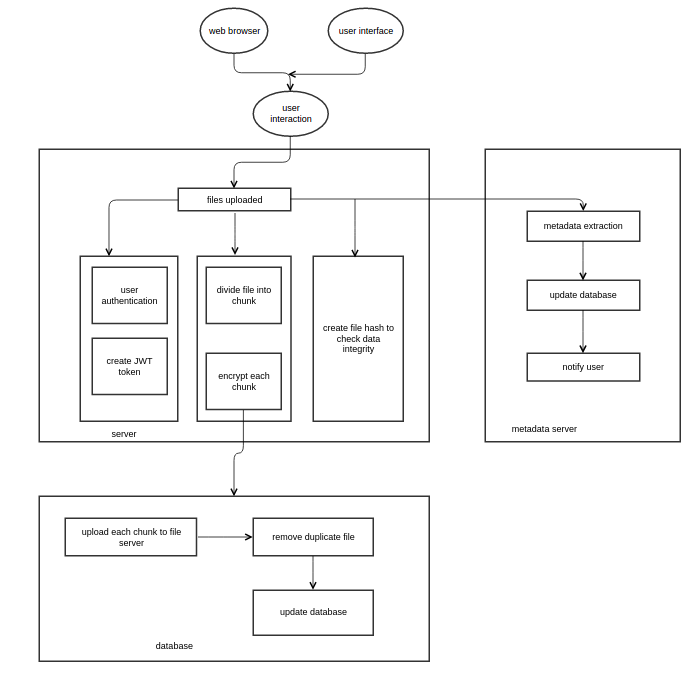


Fig 4.1 System Architecture of DOCSTASH

* 1. **System Requirements**

To achieve our objective, we need to build a platform for users to store their files. This includes hardware and software setup. The two major hardware components are a server and a storage. The server is used to process the input, maintain user data in database

**4.3.1 Hardware Setup**

Our system is setup by computer and file storage.

9

**Computer**

The computer will be used to run as a server the can interact with the user and complete user request. Servers like database server, client server, application server will run in the computer simultaneously.

**External Storage Device**

The storage devices are need to store user data that user uploads. The external storage devices need to be replicated and duplicated in order to keep the data safe.

**4.3.2 Software Setup**

**Client Server**

The client server is used to serve the static content to the user. When the user visits our website through a web browser the client server receives the request. The server responses with the static files needed for the website interface. This server can scale up and down depending upon the no of request per second and overall server load.

**API Server**

The API server is the main server that will handle all the user requests and respond to web browser.

Following are the task performed by the API server.

**• User Authentication:** Check if the user is allowed to access specific files in the server.

**• Database Updates:** It manages the database to keep the user information

updated.

**• File Storage:** It accepts the file input from the web browser and save it to the file storage.

10

1. **Design of System**
2. **System Design**

SaaS (SOFTWARE AS A SERVICE) is a software licensing and delivery model in which software is licensed on a subscription basis and is centrally hosted. It is sometimes referred to as "on-demand software”. SaaS is typically accessed by users using a thin client via a web browser.

SOFTWARE AS A SERVICE (SaaS) has gained momentum because Historically, companies were required to buy, build, and maintain their IT infrastructures despite exponential costs. SaaS gives companies an alternative. Now, they can plug in and subscribe to services built on shared infrastructure via the Internet., **DOCSTASH** is the application of the Software as a service where any user can sign up and make use of free cloud storage space.

Currently, Google Drive and One drive have developed commercial SaaS platforms to store data on cloud storage. Web applications are made user-friendly so that more users can take advantage of it. Data security is the main concern of cloud storage, the previous developments were short in providing encryption after storing files which can be exploit users data.

This project presents the development of a SaaS based web application through which users can store their important data on cloud environment without carrying it in their physical storage devices. Users will just need a smartphone or anything that has web browser with internet connectivity and they get an access to their cloud data storage.

11

**Activity Diagram:**

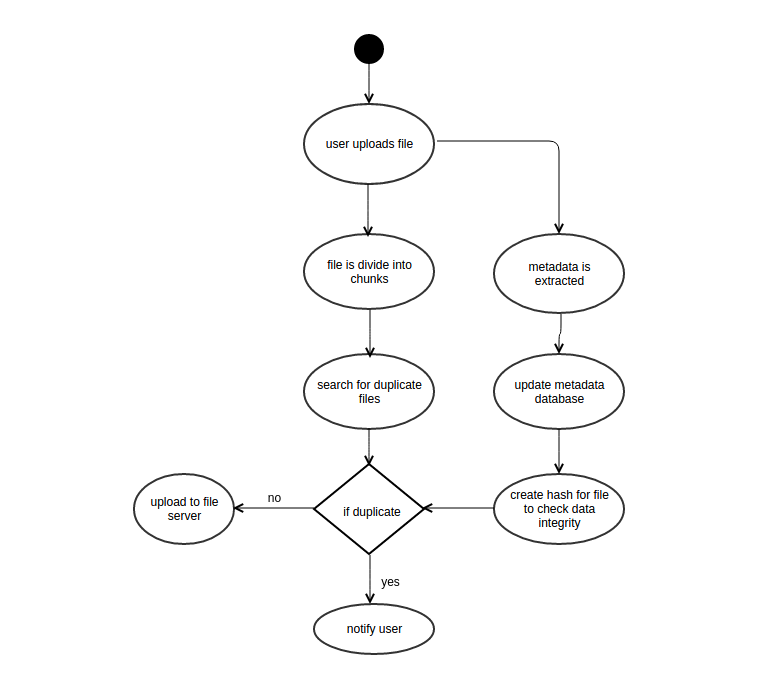
****

Fig 5.4 Activity Diagram of DOCSTASH

12

* 1. **Data Flow Diagrams**

**Level 0 Data Flow Diagrams:**

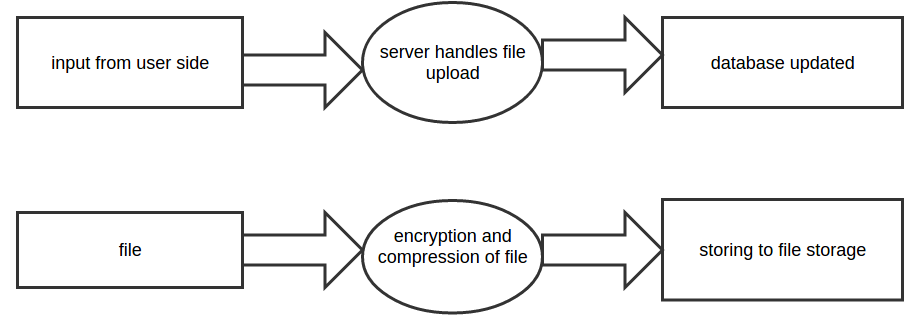


Fig 5.5 Level 0 Data Flow Diagram of User Information

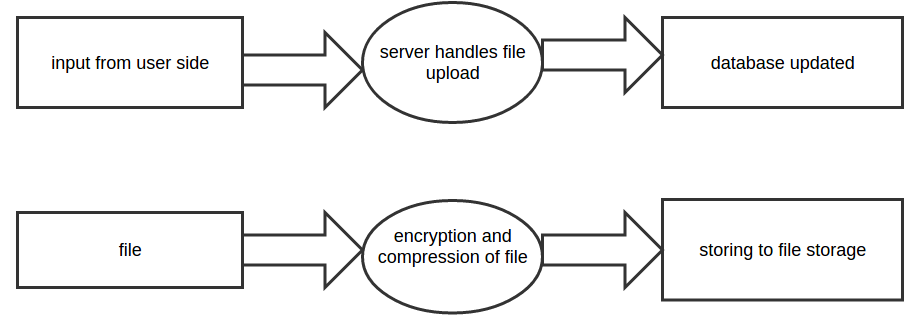


Fig 5.6 Level 0 Data Flow Diagram of File Storage

13

**Level 1 Data Flow Diagrams:**

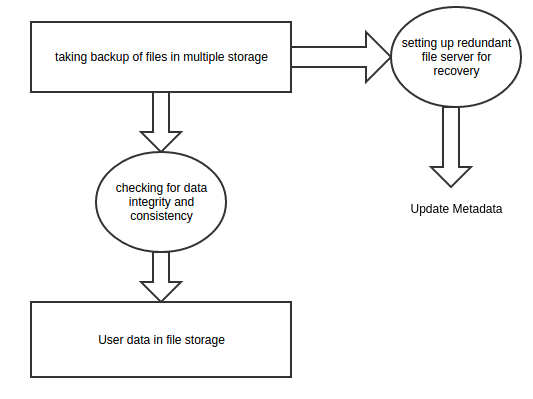
****

Fig 5.7 Level 1 Data Flow Diagram of File Storage

1. **User Interface :**

As all the interaction will be performed by the user via the web browser. The interface of the website should be very user friendly and easy to navigate. The UI is designed as a Single Page Application so that is feels like computer software as is easy to navigate as pages do not need to be refreshed regularly to get updates.

14

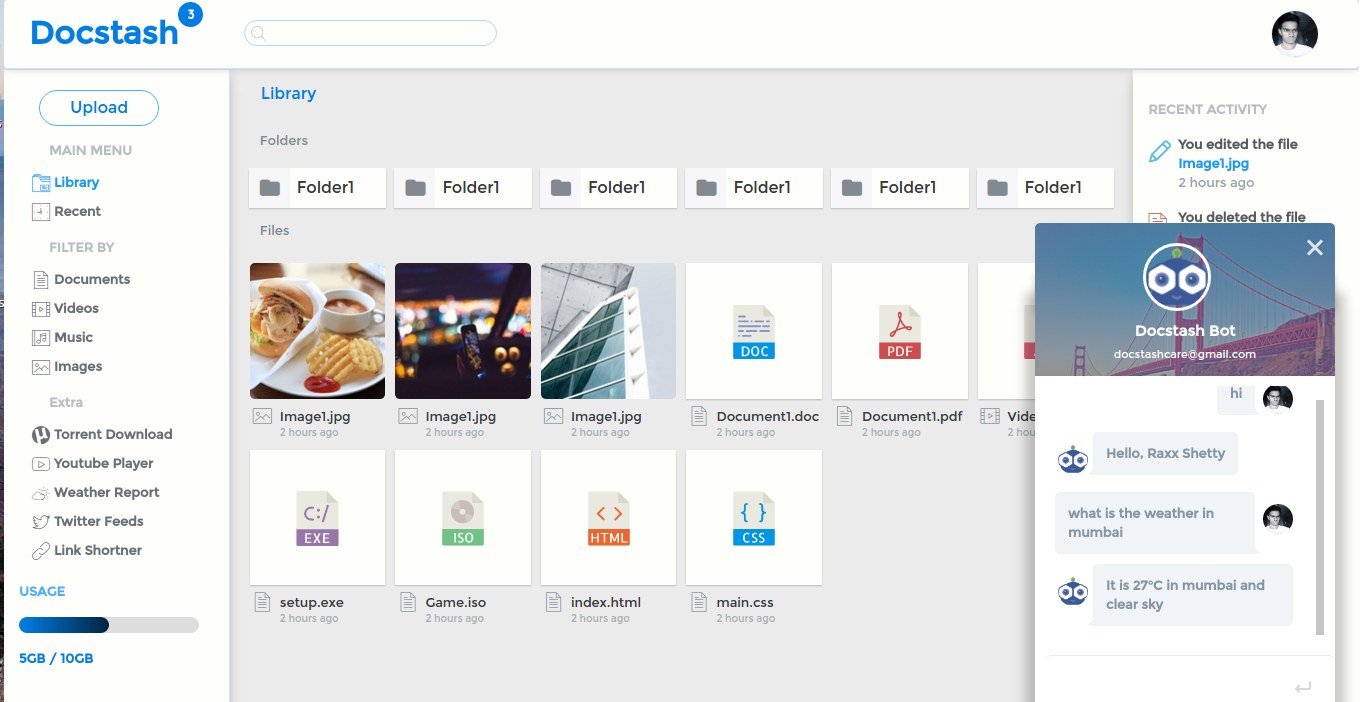


Fig 5.1 User interface

1. **API Design:**

The speed of the application depends upon the application server. It is designed as a fully REST api endpoint. This api endpoints are non-blocking and can handle multiple user request simultaneously.

User is given feedback about the status of the file uploads and downloads in real time.

1. **Database Design:**

Database schemas are designed as models and are saved in the file that can be applied to the new database easily, therefore a new database can be used very easily without need of initial configurations

1. **Development tools:**
   1. **Dev Tools:**

Using dev tools to get the current state of the of our application represented as a tree structure. This is useful to debug the application for error and see server responses in real time.

15

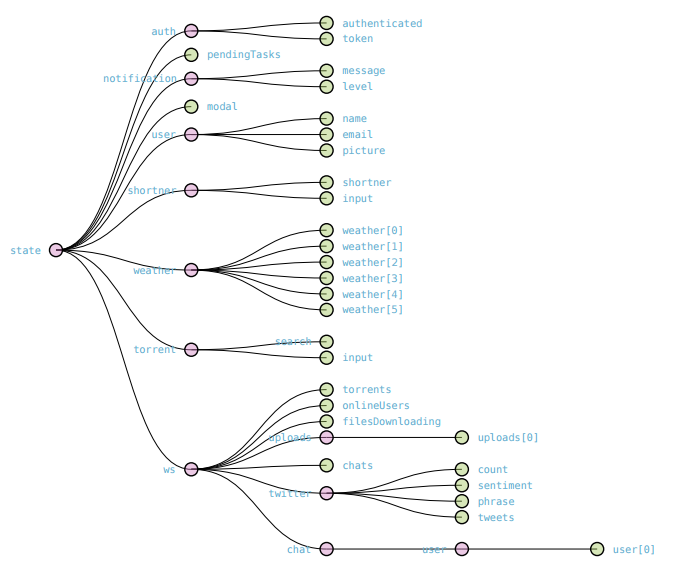
****

Fig 5.2 Dev Tool

* 1. **WEBPACK:**

Webpack is used as a module bundler, which takes modules with dependencies and generates static assets representing those modules. This is used to create a single file that can be served to the web browser that acts as a Single Page Application.

16

17

1. **Conclusion and Future Scope**
   1. **Conclusion**

Designing Cloud Storage (Docstash) is a challenging task, not least due to the evolving nature of the technology but also from the perspective of how to authenticate users with one of the best encryption technique and provide better interface experiences.

We have resolved the problem of the storage solution for any person. Docstash, which makes use of latest techniques to enhance existing features and facilities in the market provided by the good storage providers.

Docstash is a SaaS platform with increasing competitiveness that provides an enormous amount of storage through greater flexibility and optimal resource utilization to the user through authenticating itself in a very secure environment.

* 1. **Future Scope**

**Storage Enhancement:** We can use more reliable SSD (solid-state drive or solid-state disk) instead of HDD, which is currently being used.

**Enterprise Storage:** Along with personal storage, the concept can be extended to Enterprise data storage.

**Mobile Application:** Docstash can also be made available to users through mobile application for Android and iOS users.

**File Synchronization:** Automatic file sync can be implemented so that user do not have to upload files manually.

**Security:** We can use more security methods along with AES256 encryption standards to guarantee security and privacy.

17

**Literature Cited**

[1]Cloud Data Storage Technology and Its Architecture Implementation <http://www.sciencedirect.com/science/article/pii/S1877705811065192>

[2] Zero-knowledge proofs of identity Journal of Cryptology June 1988, Volume 1, Issue 2, pp 77–94 <http://link.springer.com/article/10.1007/BF02351717>

[3] CLOUD STORAGE SECURITY [http://www.ijircce.com/upload/2013/april/11\_V1204092\_O](http://www.ijircce.com/upload/2013/april/11_V1204092_O.pdf)

[4] [A\_STUDY\_ON\_CLOUD\_STORAGE](http://www.academia.edu/7193655/A_STUDY_ON_CLOUD_STORAGE_) <http://www.academia.edu/7193655/A_STUDY_ON_CLOUD_STORAGE_>

[5] 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16) [http://nec.edu.in/ICCTIDE%2716/index.html\_](http://www.academia.edu/7193655/A_STUDY_ON_CLOUD_STORAGE_)

[6] Research on Map-Reduce-based cloud storage batch auditing.

[http:// link.springer.com /ICCTIDE%2716/index.html\_](http://www.academia.edu/7193655/A_STUDY_ON_CLOUD_STORAGE_)

[7] Design and implementation of an improved cloud storage system. <http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7580922_>

[8] Concept of Cloud DRM.

https://tresorit.com/files/tresorit-drm-whitepaper.pdf

18

**Acknowledgement**

We would like to express our thanks to **Prof. Nilima Dongre**, our final year project supervisor, who has given us valuable advice and suggestions throughout the project.

We also thank our Principal **Dr. Ramesh Vasappanavara** and head of IT department**, Mr. Ashish Jadhav** for the mentoring, inspiration and support for development of the project.

Rakshan Shetty 12IT1017

Rohit Ambre 13IT3002

Nakul Chauhan 13IT1049

Shakeel Shaikh 12IT1096

19