A Mini Project Abstract

On

DATA DUPLICATION REMOVAL TECHNOLOGY USING AWS SERVICES

Submitted to **CMR ENGINEERING COLLEGE**

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted By

MOSARLANITHIN REDDY (208R1A0536)

SOMYA TRIPATHI (208R1A0554)

ANAGANDULA RAJITHA (218R5A0501)

JUPALLI JEEVAN KUMAR (208R1A0525)

Under the guidance of

Mrs. M.Prashanthi

Assistant Professor, Department of CSE

Department of Computer Science & Engineering

CMR ENGINEERING COLLEGE

UGC AUTONOMOUS

(Accredited by NBA Approved by AICTE,NEW DELHI, Affiliated to JNTU Hyderabad)

Kandlakoya, Medchal Road, Medchal Malkajgiri Dist. Hyderabad-501 401)

2023-2024

ABSTRACT

Cloud computing comes all through core interest advancement of network computing, virtualization just as web advancements. With an expansion in the use of cloud storage, powerful strategies should be utilized to diminish equipment costs, meet the data transmission necessities, and build stockpiling proficiency.

EXISTING SYSTEM

This can be accomplished by utilizing Data Deduplication. Using this, less information will be on the server thus require less equipment and users would have the option to put more data in the additional space available. At present utilization of cloud storage is expanding and to conquer expanding information issues, Data Deduplication methods are utilized. Information Deduplication methods can’t be applied straightforwardly with security instruments.

PROPOSED SYSTEM

In this paper, we are eliminating copy information to save storage space and speed up the organization. Here we applied MD5 hashing to generate a hash value (when uploaded to cloud storage) and then compare those with values (when the same file uploaded with a different name) to find out the duplicate data in the cloud environment. When deduplication is accomplished framework plan for secure information change in the organization. Security is accomplished through encryption and decoding of the information. This paper inspects secure deduplication strategy. After removal of duplicate data pointers will give reference to the original file.

HARDWARE REQUIREMENTS

MINIMUM (Required for Execution)		MY SYSTEM (Development)
System	Pentium IV 2.2 GHz	i3 Processor 5th Gen
Hard Disk	20 Gb	500 Gb
Ram	1 Gb	4 Gb

SOFTWARE REQUIREMENTS

Operating System	Windows 10/11
Development Software	Python 3.10
Programming Language	Python
Integrated Development Environment (IDE)	Visual Studio Code
Front End Technologies	HTML5, CSS3, Java Script
Back End Technologies or Framework	Django
Database Language	SQL
Database (RDBMS)	MySQL
Database Software	WAMP or XAMPP Server
Web Server or Deployment Server	Django Application Development Server
Design/Modelling	Rational Rose

Mini Project Coordinator **Internal Guide**