Implementation: For implementation we preferred Java language, IntelliJ IDEA CE framework and Windows O.S. Platform as it provides inbuilt server called IIS. Java provides inbuilt *MessageDigest* class for SHA-256 hashing. The SHA (Secure Hash Algorithm) is one of the popular cryptographic hash functions. A cryptographic hash can be used to make a signature for a text or a data file. Data deduplication is referred to as a strategy offered to cloud storage providers (CSPs) to eliminate the duplicate data and keep only a single unique copy of it for storage space saving purpose.

Data deduplication is one of the techniques which used to solve the repetition of data. The deduplication techniques are generally used in the cloud server for reducing the space of the server. To prevent the unauthorized use of data accessing and create duplicate data on cloud the encryption technique to encrypt the data before stored on cloud server. Cloud Storage usually contains business-critical data and processes; hence high security is the only solution to retain strong trust relationship between the cloud users and cloud service providers

In this methodology we have to detect the duplicate copy of the file any type of file can be detect file .txt,.doc,.xls, ppt, .pdf. so we have to start with uploading the file when we upload the file we have to extract first 50 bytes from the file and last 50 bytes from the file. After extracting this 100 byte we have match this byte with existing byte. This comparing is done one by one byte i.e.one byte after another upon last byte. After completion of comparing this byte if this new file upload file is duplicate then we will discard the file. If this file is not duplicate, then we will upload this file. In this way the methodology is use.

