

## **Application Aware Bid Data Deduplication in cloud Environment**

### **Abstract:**

Many organizations and individual persons are showing interest to store data in the cloud because of the advantages of the data storage in the cloud. So along with these advantages security is the more important concern taken to be care of. When uploading the data it will be converted into encrypted format so that there is a chance for uploading the duplicate data this is another thing to taken care about data storage.

### **Existing system:**

There are several existing solutions that aim to tackle the above two challenges of distributed deduplication by exploiting data similarity or locality. Locality means that the chunks of a data stream will appear in approximately the same order again with a high probability. Locality-only based approaches distribute data across deduplication servers at coarse granularity to achieve scalable deduplication throughput across the nodes by exploiting locality in data streams, but they suffer low duplicate elimination ratio due to high cross-node redundancy.

### **Disadvantages:**

1. Finding the data deduplication among several data centers.
2. Data has to be distributed on multiple locations.

### **Proposed system:**

We propose AppDedupe, an application-aware scalable inline distributed deduplication framework in cloud environment, to meet this challenge by exploiting application awareness,

**Further Details Contact: A Vinay 9030333433, 08772261612, 9014123891**  
**#301, 303 & 304, 3rd Floor, AVR Buildings, Opp to SV Music College, Balaji Colony, Tirupati - 515702**  
**Email: info@takeoffprojects.com | www.takeoffprojects.com**

data similarity and locality to optimize distributed deduplication with inter-node two-tiered data routing and intra-node application-aware deduplication. It first dispenses application data at file level with an application-aware routing to keep application locality, then assigns similar application data to the same storage node at the super-chunk granularity using a hand printing-based stateful data routing scheme to maintain high global deduplication efficiency, meanwhile balances the workload across nodes.

### **Advantages:**

1. Stores the similar type of data in a single location.
2. so that data duplication will be found out easily.

### **SYSTEM REQUIREMENTS**

#### **H/W System Configuration:-**

- Processor - Pentium –III
- RAM - 256 MB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

#### **S/W System Configuration:-**

- Operating System : Windows95/98/2000/XP
- Application Server : Tomcat5.0/6.X
- Front End : HTML, Jsp
- Scripts : JavaScript.
- Server side Script : Java Server Pages.
- Database : MySQL 5.0

- Database Connectivity : JDBC

[www.takeoffprojects.com](http://www.takeoffprojects.com)