SAPTHAGIRI COLLEGE OF ENGINEERING

14/5, Chikkasandra, Hesaraghatta Main Road, Bangalore-560057

Department of Computer Science and Engineering

Certificate

Certified that the project work entitled "**An optimized approach for accessing small files in hd using TLB Mapfile**" carried out by **Rakesha R S(1SG13CS082), Akshay K(1SG14CS401), Annap V(1SG14CS403), Rohith Murthy S K(1SG14CS418)**, bonafide students of **Sapthagiri College Engineering**, in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science a Engineering** of **Visvesvaraya Technological University, Belgaum** during the academic year **2016-17.** It certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the rep deposited in the department library. The project report has been approved as it satisfies the academic requireme in respect of Project work (10CS85) prescribed for the said degree.

Signature of the Guide	**Signature of the HOD**	**Signature of the Principal**
Prof. Sushmitha J	**Dr. Prashanth C.M**	**Dr. Aswatha Kumar M**
Assistant Professor	Professor & Head	Principal

Dr. Aswatha Kumar. M
Principal
Sapthagiri College of Engineering
No. 14/5, Chikkasandra,
Hesaraghtta Main Road,
Bangalore-560 057

Name of the Examiners	**Signature with date**
1.	1. .....
2	2. .....

ABSTRACT

This project presents an access optimization approach for HDFS small file based on MapFile: TLB-MapFile. TLB-MapFile merges massive small files into large files by MapFile mechanism to reduce NameNode memory consumption and add fast table structure (TLB) in DataNode, and to improve retrieval efficiency of small files. First, according to MapFile mechanism, small files are merged into large files and stored in HDFS. Second, the access frequency and the ordered queue of small files (per unit time) can be obtained through accessing system logs in HDFS, and the mapping information between block and small files are stored in the TLB table with regularly being updated. Experiment results show that this method can effectively reduce NameNode memory consumption and improve the accessing speed of small files.