

Available Online at [www.ijcsmc.com](http://www.ijcsmc.com)

## International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

*IJCSMC, Vol. 3, Issue. 2, February 2014, pg.463 – 466*

### **REVIEW ARTICLE**

# Privacy in Map Reduce Based Systems: A Review

Rosmy C Jose, Shaiju Paul

Department of Computer Science & Engineering, Jyothi Engineering College, Cheruthuruthy, Kerala, India  
[rosmycJose29@gmail.com](mailto:rosmycJose29@gmail.com), [intouch.paulson@gmail.com](mailto:intouch.paulson@gmail.com)



**Abstract**—Today, every organisation generates and adds huge amount of data to the cloud. This vast amount of data which cannot be effectively captured, processed and analysed by traditional database and search tools is called Big Data. The processing of big data is made possible by using Map Reduce, a programming model and an associated implementation, introduced by Google. MapReduce process data, which are located at different data nodes. It pushes computations to where the data resides rather than the opposite. So, Map Reduce Framework or source codes may leak sensitive data during computation process. In current implementation (Airavat) Mapper code is written by user and Reducer code is selected from a list provided by the system. If these codes are given by the system itself, usability may become low. Therefore, in the proposed system both Map and Reduce codes can be written by the user. So usability will be high. A Computation System ensures the privacy leak through storage channels (network connections, files) or privacy leak through the output of the computation is stopped. Use SELinux is used to prevent storage channel leaks. Leaks through the output of the computations are checked by using differential privacy mechanisms.

**Keywords**- BigData, MapReduce, Hadoop, Airavat

Full Text: <http://www.ijcsmc.com/docs/papers/February2014/V3I2201475.pdf>