**Detecting Suspicious File Migration or Replication in the Cloud**

**Abstract:**

There has been a prolific rise in the popularity of cloud storage in recent years. While cloud storage offers many advantages such as flexibility and convenience, users are typically unable to tell or control the actual locations of their data. This limitation may affect users' confidence and trust in the storage provider, or even render cloud unsuitable for storing data with strict location requirements. To address this issue, we propose a system called LAST-HDFS which integrates Location-Aware Storage Technique (LAST) into the open source Hadoop Distributed File System (HDFS). The LAST-HDFS system enforces location-aware file allocations and continuously monitors file transfers to detect potentially illegal transfers in the cloud. Illegal transfers here refer to attempts to move sensitive data outside the (“legal”) boundaries specified by the file owner and its policies. Our underlying algorithms model file transfers among nodes as a weighted graph, and maximize the probability of storing data items of similar privacy preferences in the same region. We equip each cloud node with a socket monitor that is capable of monitoring the real-time communication among cloud nodes. Based on the real-time data transfer information captured by the socket monitors, our system calculates the probability of a given transfer to be illegal. We have implemented our proposed framework and carried out an extensive experimental evaluation in a large-scale real cloud environment to demonstrate the effectiveness and efficiency of our proposed system.