

## Improving Accessing Efficiency of Cloud Storage Using De-Duplication and Feedback Schemes

### Abstract:

File distribution and storage in a cloud storage environment is usually handled by storage device providers or physical storage devices rented from third parties. Files can be integrated into useful resources that users are then able to access via centralized management and virtualization. Nevertheless, when the number of files continues to increase, the condition of every storage node cannot be guaranteed by the manager. High volumes of files will result in wasted hardware resources, increased control complexity of the data center, and a less efficient cloud storage system. Therefore, in order to reduce workloads due to duplicate files, we propose the index name servers (INS) to manage not only file storage, data de-duplication, optimized node selection, and server load balancing, but also file compression, chunk matching, real-time feedback control, IP information, and busy level index monitoring. To manage and optimize the storage nodes based on the client-side transmission status by our proposed INS, all nodes must elicit optimal performance and offer suitable resources to clients. In this way, not only can the performance of the storage system be improved, but the files can also be reasonably distributed, decreasing the workload of the storage nodes.