**Abstract**

Cloud storage services, such as Dropbox, Microsoft SkyDrive and Google Drive, have gained enormous popularity in recent years. They offer users with convenient and reliable approaches to store and share data from anywhere, on any device, at anytime. However, they are suffering from the “traffic overuse problem” in the presence of frequent, short data updates [1][2]. To address this problem, we are implementing a tunable cloud storage service (named “T-CloudDisk” or “ThuCloudDisk”) for flexible batched synchronization. This poster briefly introduces its characteristics, technical approach, and preliminary implementation.

**Traffic overuse problem.** Saving network traffic is a critical goal of cloud storage services, given that billions of files are synchronized to the cloud every day. To minimize the network overhead, cloud storage services (e.g., Dropbox) employ binary diff, data compression, and other mechanisms when delivering updates among users. However, despite these optimizations, we observe that in the presence of frequent, short updates to user data, the network traffic generated by cloud storage services often exhibits pathological inefficiencies.

Through comprehensive measurements and detailed analysis, we demonstrate that many cloud storage applications generate session maintenance traffic that far exceeds the useful update traffic, as shown in Fig.1. We refer to this behavior as the “traffic overuse problem” [1][2].

**UDS middleware.** To address this problem, we propose the UDS middleware. UDS batches updates from clients to significantly reduce the overhead caused by session maintenance traffic, as depicted in Fig.2.

Although UDS significantly reduces the traffic overuse, acting as a middleware between the user’s local filesystem and the concerned cloud storage application (like Dropbox), it requires considerable additional storage space in the user’s local disk, and currently it uses fixed buffer size and timer threshold for batched sync.

**The T-CloudDisk (ThuCloudDisk) System Design and Implementation**

**T-CloudDisk.** In order to further overcome the shortcomings of UDS, we are implementing an independent, tunable cloud storage service (named “T-CloudDisk” or “ThuCloudDisk”) for flexible batched synchronization:

- **Independent:** T-CloudDisk is a self-contained cloud storage service, rather than a middleware that relies on other services’ apps. As depicted in Fig.3, it has its own user app and private cloud that maintains the critical meta-data and user liveness information. With such an independent service, we are able to conduct more in-depth, white-box research on cloud storage.

- **Tunable:** To avoid the expensive infrastructure cost and enhance the service scalability, T-CloudDisk outsources all file contents to a public cloud that supports RESTful APIs, such as Amazon S3, Aliyun.com OSS, and Openstack Swift (our current implementation).

- **Flexible:** The T-CloudDisk app allows users to customize the buffer size and timer threshold according to their specific requirements, as shown in Fig.4. This increases the flexibility of batched sync as it enables users’ own tradeoff between traffic usage and user experience.

**Related Work**


**Contact**

* Zhenhua Li, School of Software and TNLIST, Tsinghua University, Beijing, China. lizhenhua1983@tsinghua.edu.cn. Homepage: http://www.greenorbs.org/people/lzh.

* Linsong Cheng, School of Software, Tsinghua University, Beijing, China. chenglsl10@mails.tsinghua.edu.cn.