Shortening of Transfer Duration Using File Transfer Striping

Hidekatsu Kayama, Toshiaki Hishinuma, and Eijiroh Ohki

Faculty of Informatics, Kogakuin University, Tokyo, Japan
E-mail: j209038@ns.kogakuin.ac.jp

Abstract

One of problems in the online storage systems is shortening of transfer duration of big files. In order to shorten transfer duration, it can be thought that a transfer method is effective: a system divides files to any block length, after the system parallel transfers the blocks to servers; this method is named "FTS (File Transfer Striping)". In this research, so as to verify F.T.S. is effective for shortening transfer duration, we implemented an online storage system leveraging F.T.S. employing FTP, and conducted a measure experiment of transfer duration of the system. From the result of the experiment, it has the potential to shorten duration when uploading. On the other hand, a transfer rate in downloading remains an issue, it is necessary to analyze bottlenecks of the system.

Online storage service has been continuingly expanding several years. The large storage service provides to over 500,000 users [1]; moreover the size of the domestic market is estimated more than 26 billion yen, and foresees that it continues to grow [2].

One of the problems in the online storage systems is transfer duration of big file. If one server is used, a transfer rate is shorter than a band limitation of the server, therefore the transfer duration is increased in proportion to the file size.

I focused on the striping (RAID-0) method, which makes access of several hard disks faster. It can be thought that increasing number of servers and uploading files parallel is possible to shorten of file transfer duration. This method is named "F.T.S. (File Transfer Striping)". The overview of F.T.S. is as Fig1.

In this research, I intend to verify effectiveness of F.T.S. for shortening transfer duration. In order to verifying that F.T.S. is effective for shortening transfer duration, I implemented an online storage system leveraging F.T.S. and conducted measure experiments of transfer duration of the system.

This system developed in the environment of Table1; moreover FTP (File Transfer Protocol) is used as a transmission mode of files. Although FTP had already become an old technology [3], the system adopts FTP because it is comparatively ubiquitous to be implemented and compared speed.

In the implemented system, client application uploads and downloads files. Major processes without authentication of FTP are done in client, therefore setting of servers are simple. A process of dividing and storing a file is as Fig2.

Table 1. Environment of Development

Development language	C++
Compiler	Microsoft Visual Studio 2010
Libraries	WinInet, Boost C++ Libraries

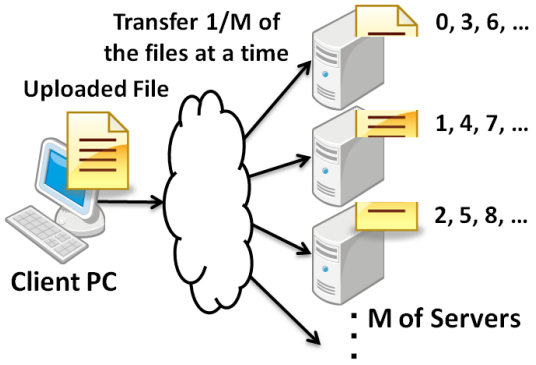

Fig. 1. The Overview of the File Transfer Striping

In order to verify effectiveness of FTS to shorten transfer duration, we conducted a measure experiment of file transfer duration of the system as Fig3. The application and other files deployed on the RAM disk. The capabilities of each device are as Table2. The experiment condition is as Table3.

In measurement of upload time, we deployed files have been divided on secondary storage in client PC preliminarily, and measured the time between start of the transmission and completion of the transmission of all division files.

Also in measurement of download time, we deployed block files on servers preliminarily, and measured the time between start of the loading a first block file and completion of writing of all division files.

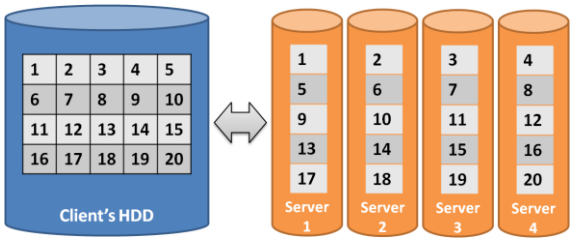

Fig. 2. The System of Dividing and Storing of a File

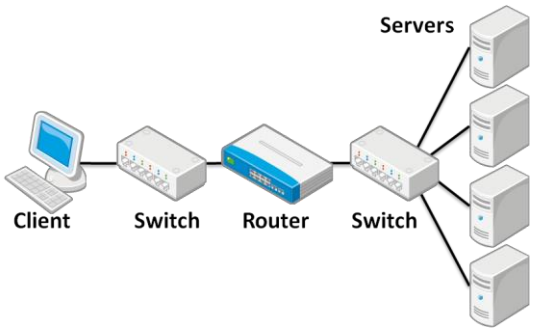

Fig. 3. Environment of the Experiment (The File Transfer Striping Using FTP)

Table 2. Environment of the Experiment

PC		
	OS	Windows7 Enterprise (64bit)
	CPU	Intel Core i7 2600 3.4GHz
	Memory	12GB
	RAM Disk	BUFFALO RAMDISK Utility Ver. 3.0.0.0.
	NIC	1Gbps
Switch		BUFFALO LSW4-GT-8EP/WH 1000BASE-T
Router		I-O DATA ETG2-R 1000BASE-T
Server		I-O DATA NAS HDL-A (2.0TB HDD)

Table 3. Condition of the Experiment

Size of the file to transmit (MB)	400, 800, 1024, 2048
Block length	8MB
The number of servers	4 sets
The number of times of trial of each conditions	30 times

For a comparison purpose, we measured time of transmitting a file to one of server using single FTP. The system environment is as Fig 4.

As a result of this experiment as shown in Fig5, I reached a conclusion that the F.T.S. has a possibility that make transfer rate of files to be about 2.1 times higher than traditional FTP system at a maximum when uploading. It can be thought that the transfer rates of F.T.S relate size of translation files, therefore the rates have possibility to be increased when transferring larger files.

On other hand, in download, the transfer rates of F.T.S. are lower than traditional FTP system. It was expected that the rates in downloading are close to in uploading, although the result shows different.

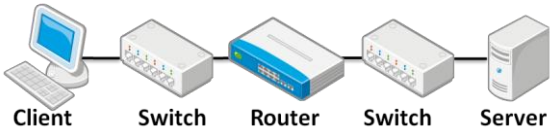

Fig. 4. Environment of the Experiment (Single FTP)

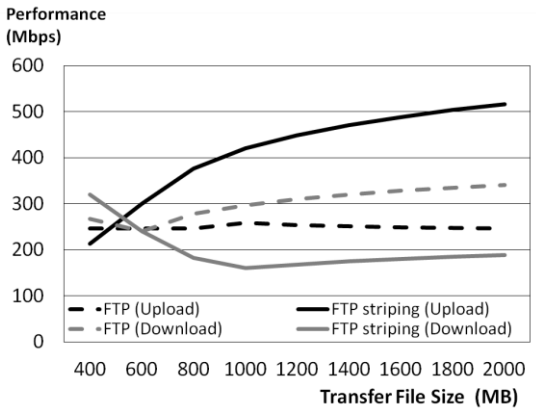

Fig.5 . Comparing of Transfer Rate

It can be thought about the reason that the times measured in this experience include accessing of file system, therefore differences between read and write become bottlenecks. In FTP, a completion of file translation is when the data has written on receiver as a file, therefore it happens access of file system at every transfer. Therefore, it has necessary to measure the "pure" transfer durations when uploading and downloading, and thereafter compare them.

As a result of this research, F.T.S. has possibility to shorten transfer duration when uploading. On the other hand, its effectiveness when downloading should be researched more, and performance could be improved, I believe. Additionally, I consider another experiment in the real environment conditions, such as a number of servers, clients, line speed, and etc.

References:

[1] "The number of month-long users of cloud service Dropbox and Evernote break through 500,000 people", The Nielsen Company Japan, http://www.netratings.co.jp/, Accessed 2012/8/18.

[2] "Venture business clairvoyance: The difference between "Storage as a Service", "Cloud storage", and "online storage"", Tadashi Yoshimasa, ITmedia Inc. , http://blogs.itmedia.co.jp/yoshimasa/2012/05/storage-as-a-se-be25.html, Accessed 2012/8/18.

[3] "Serialization: Explication of internet protocol (10), FTP (File Transfer Protocol) ~ first part", Kaiji Yanase, ITmedia Inc. , http://www.atmarkit.co.jp/fnetwork/rensai/netpro10/netpro01.html, Accessed 2012/9/11.