Research on Analysis of Factors Influencing Terrestrial Transmission Efficiency of Satellite Data Files

Zhang Jing
Institute of Remote Sensing and Digital Earth
Chinese Academy of Sciences
Beijing, China
zhangjingswcd@126.com

JDIM
Journal of Digital Information Management

ABSTRACT: The paper firstly proposes two factors influencing transmission rate of satellite data file, the number of files and threads. The paper uses the control variate method to design the test scheme of the transmitting satellite data file, and changes the values of variables for lots of repeated tests to achieve the mean transmission time. And the paper compared the file transmission rate under three different variables, which draws the conclusion that the transmission efficiency of using single-file and multi-thread variable combination is the best. And the paper makes deep study on the selection of the number of threads. In the test environment of network delay of 50ms, the transmission time of using 3 or 4 threads to transmit data files is short, which proves that the number of threads transmitting files needs to be determined by the network environment.

Subject Categories and Descriptors
H.4.3 [Communications Applications]: B.4.1 Data Communications Devices]: Transmitters

General Terms: Satellite Communication, Transmission Rate

Keywords: Satellite Data, Transmission, Number of Files, Number of Threads, Transmission Time, Transmission Efficiency

Received: 11 November 2013, Revised 13 February 2014, Accepted 10 February 2014

1. Introduction

In recent years, remote sensing satellite develops rapidly and has expanded from military scientific research field to civilian field. Remote sensing satellite can provide lots of ground images with greater accuracy, which has great reference value of scientific research and navigation. But the remote sensing satellite has greater data volume. And it has high performance requirements on satellite-ground transmission and ground transmission phase. With the increase of the number of civilian remote sensing satellites, the performance requirements are promoted. The limitation of the geographic position of satellite data receiving station makes it needs to receive pooling analysis of wide area network transmitting to ground station center after receiving satellite data. The programs depended by the paper include three satellite data receiving stations which are distributed in the south, west and north of China. And there are a satellite data receiving station locating in the ground station center of Beijing. Three satellite data receiving stations need to transmit the data to the ground station center in Beijing after receiving data.

For the hardware, adding network acceleration devices is generally used at home and abroad. The essence is to deploy hardware, network accelerator on satellite data receiving station, to use the data compression function of network accelerator to compress the transmitted data files, and then to transmit the data files, which can improve the transmission rate of satellite data files. But it still can't satisfy the increasing transmission requirements of

satellite data. And different network accelerators play different roles in improving the transmission rate of satellite data files, which depends on the compression algorithm used by the network accelerator and the invested cost. For business software, the overseas and domestic research focus on the optimization of TCP protocol, the selection of RTT value [1] and flow control, and has got achievements. But there are few researches about the performance of macro data transmission and the influence of transmission schemes on transmission performance. So the paper focuses on studying the selection of macro data satellite data transmission schemes, makes detailed analysis and gets the final research results.

TCP protocol has good performance on Ethernet. And the optimized TCP can satisfy higher performance requirements. But the other methods about performance, promotion still need to be considered from the other aspects [2]. On one hand, the characteristics of satellite link are that the round-trip time is greater [3] (> 50ms) , and the link capacity is smaller (the greater the capacity is, the higher the rental cost). On the other hand, TCP is a mature protocol, and optimizing it can't improve the transmission performance greatly [4].

The paper starts from the research on application layer, proposes that the number of threads and files may influence the transmission efficiency of satellite data, and makes comparative analysis on different transmission methods by using control variate method [5]. The final objective is to combine the experimental data to make comprehensive comparative analysis and find out the scheme with the greatest transmission efficiency.

2. Research Methods Design

The paper designs three different transmission schemes for study, as follows.

a) Single-file and single-thread transmission.

b) Single-file and multi-thread transmission.

c) Multi-file and multi-thread transmission.

Single-thread data file transmission only has and uses a thread transmission data. Multi-thread file transmission means to transmit a data file or several data files in parallel at the same time.

In the process of experiment, the test program is simulated to make ground transmission of satellite data file. The network used for transmitting data file is delayed WAN. The transmission protocol is TCP, and the protocol is not optimized [6]. Two real servers are selected as the server and client side of transmitting files.

(1) The premise of comparing the above three data file transmission modes is to transmit the file with the same data size. And the time for each scheme to transmit data files is compared. The third scheme can transmit several data files, but the total data size of the transmitted files is the same to that of the other two modes.

(2) The test environment is relatively independent, and there is no the disturbance of the other transmission data.

(3) The read-write rate of the hard disc of the server is greater than the greatest transmission rate of network bandwidth.

(4) It is general that single-file and multi-thread transmission scheme can use two threads, and multi-file and multi-thread transmission uses two files for transmission at the same time. And each file can be transmitted by a thread.

(5) The data volume of transmitting files should be large, or the early end of transmission can have great influence on the results. The data volume in the paper is 2G.

(6) Each scheme receives several times of satellite data file transmission. And the scheme using the shortest time to transmit data files is achieved by comparing the average transmission time of different schemes of transmitting satellite data files.

3. Analysis on Influencing Factors of File Transmission Efficiency

3.1 Experiment environment

In order to ensure that test data can accurately reflect the transmission efficiency of real files for each scheme, the experiment is made in the real wide area network. The assigned data files are transmitted from the data receiving station server of the south to the ground station center server in Beijing. And in the experiment, there are no other applications occupying network bandwidth resources in the network. For convenience, the following is an explanation of the key content of the experiment.

Server: There are two servers transmitting data files, server A and server B. Two servers are equipped with network interface card. Two servers are connected by Gigabit switch. And RHEL5.5 operation system is installed.

ServerSS: It is the name of executable program and file-transmission server (transmitter) program. It is deployed on server A, and uses single thread to send a file. The paper uses control variate method to make lots experiment on three variable combinations, and then the achieved data is compared and analyzed. Lastly, the best variable combination is achieved. The detailed explanation and experiment on each mode are as follows.

ClientSS: It is the name of executable program and file-transmission server (transmitter) program. It is deployed on server B, and uses single thread to receive a file.

ServerSM: It is the name of executable program and file-transmission server (transmitter) program. It is deployed

on server A, and uses multi threads to send a file. And each thread sends a part of data.

ClientSM: It is the name of the executable program and file-transmission server (transmitter) program. It is deployed on server B, and uses multi threads to receive a file. And each thread receives a part of data.

ServerMM: It is the name of executable program and file-transmission server (transmitter) program. It is deployed on server A, and uses multi threads to send multi files at the same time.

ClientMM: It is the name of executable program and file-transmission server (transmitter) program. It is deployed on server B, and uses multi threads to receive multi files.

2GFile: It is a data file. The size is 2G, and it is deployed on the server A.

1GFile1 and 1GFile2: It is a data file. The size is 1G, and it is deployed on the server A.

3.2 Experiment procedure

It is simulated that the specified quantity of data files is sent from the server A, and the total size of data files for each experiment is the same. The time of every transmission is recorded, and the average transmission time of each scheme is solved, which can get the scheme using the shortest time to transmit data files.

As the experimental procedures of each scheme are nearly the same, each scheme is tested by steps, and there is little difference for applications and the selection of data files, the experiment procedures are as follows.

1. The transmitter program of server A is started. The terminal prompts to input the name of the transmitted file. The name of a file or multi files is input into the system terminal, and multi filenames are separated by spaces.

2. The receiving end program on server B is started to establish connection for receiving data files automatically.

3. After transmitting the files, the server A terminal shows the transmission time for record.

4. The experiment is repeated for 15 times, which can get the mean value.

5. The experiment is over, and the average time of the scheme transmitting data files is computed.

3.3 Experiment procedure

The selected variables are different, so they are described respectively in the following three chapters. And each chapter represents a variable combination mode.

3.3.1 Selected variables:single file and single thread

Single-file and single-thread file transmission means that the simulation test program only can establish a thread to transmit a data file. The size of the transmitted file is 2G. The simulation test program transmits the data file from server A to server B. After the transmission, the server A records the time of sending data, and the bandwidth utilization ratio is computed. The experimental results are shown in Table 1.

The simulation test program and data files used for the experiment are as follows.

Simulation test program: The transmitter program is ServerSS and is deployed on server A. The receiving-end program is ClientSS and is deployed on server B. SS of ServerSS and ClientSS means single file and single thread.
Data file: 2GFile

Experiment number	Send file	Time /s
1	2GFile	29.04
2	2GFile	27.80
3	2GFile	26.01
4	2GFile	26.77
5	2GFile	26.57
6	2GFile	27.50
7	2GFile	28.41
8	2GFile	28.39
9	2GFile	28.91
10	2GFile	26.70
11	2GFile	26.09
12	2GFile	28.40
13	2GFile	27.85
14	2GFile	27.38
15	2GFile	27.55
Average		27.56

Table 1. Testing time of single-file and single-thread transmission

We can see from the experiment results in Table 1 that the average time of single file and single thread transmitting data files is 27.56s, and the bandwidth utilization is 59.45%.

3.3.2 Selected variables: single file and multi threads

Single-file and multi-thread file transmission means the simulation test program establishes multi threads to transmit a single data file, and the size of the file is 2G. As two threads are used in the experiment, and each thread reads different parts of the file, each thread transmits the first part of data and the latter part of data for the transmitted file. The simulation test program sends the data file from the server A to server B. After transmission, the server A records the time to transmit data, and the bandwidth utilization is computed. The experimental results are shown in Table 2.

The simulation test program and data files used for the

experiment are as follows.

Simulation test program: The transmitter program is ServerSS and is deployed on server A. The receiving-end program is ClientSM and is deployed on server B. SM of ServerSM and ClientSM means single file and single thread.

Data file: 2GFile.

Experiment number	Send file	Time /s
1	2GFile	20.01
2	2GFile	22.26
3	2GFile	20.02
4	2GFile	20.47
5	2GFile	21.17
6	2GFile	21.21
7	2GFile	20.30
8	2GFile	19.79
9	2GFile	20.80
10	2GFile	20.18
11	2GFile	19.76
12	2GFile	20.42
13	2GFile	19.95
14	2GFile	20.55
15	2GFile	20.32
Average		20.48

Table 2. Testing time of single-file and multi-thread transmission

We can see from the experiment results in Table 2 that the average time of single file and multi thread transmitting data files is 20.48s, and the bandwidth utilization is 80%.

3.3.3 Selected variables: multi files and multi threads

Multi-file and multi-thread file transmission means that the simulation test program can establish multi threads to transmit multi data files. In the scheme, we firstly transmits two data files, and the size of the files is 1G. And the total data volume is 2G. Each thread takes charge of the transmission of a data. The simulation test program sends the data file from the server A to server B. After sending two data files, the server A records the time of sending data, and the bandwidth utilization is computed. The experimental results are shown in Table 3. The simulation test program and data files used for the experiment are as follows.

Simulation test program: the transmitter program is ServerM and is deployed on server A. The receiving-end program is ClientM and is deployed on server B. M of ServerSM and ClientM means multi files and multi threads.

Data files: 1GFile1, 1GFile2.

Experiment number	Send file	Time /s
1	1GFile101GFile2	23.51
2	1GFile101GFile2	23.47
3	1GFile101GFile2	22.48
4	1GFile101GFile2	23.39
5	1GFile101GFile2	23.34
6	1GFile101GFile2	23.43
7	1GFile101GFile2	23.51
8	1GFile101GFile2	23.14
9	1GFile101GFile2	23.37
10	1GFile101GFile2	24.17
11	1GFile101GFile2	23.39
12	1GFile101GFile2	23.08
13	1GFile101GFile2	22.89
14	1GFile101GFile2	23.71
15	1GFile101GFile2	22.73
Average		23.31

Table 3. Testing time of multi-file and multi-thread transmission

We can see from the experiment results in Table 3 that the average time of single file and multi thread transmitting data files is 23.31s, and the bandwidth utilization is 70.23%.

3.4 Experimental data analysis

The paper analyzes the transmission result data of three file transmission modes. Under the premise of transmitting the data files with the same data amount, the paper compares the advantages and disadvantages of the transmission time of three file transmission modes, as shown in Figure 1.

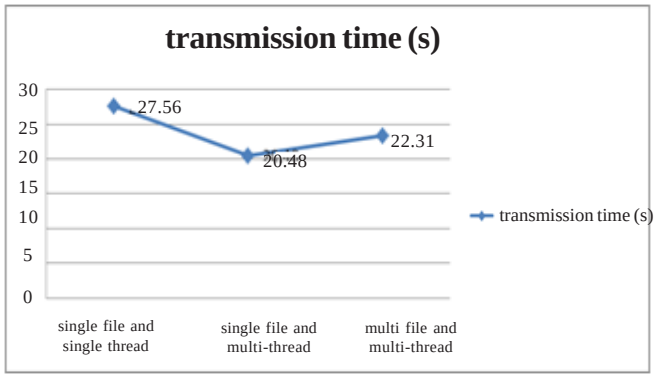

Figure 1. Comparison on transmission time of each scheme

Figure 1 is the average transmission time of lots of repeated experiments that three different variable combinations transmit data files. From figure 1, we can see that single-file and multi-thread transmission can complete the transmission with the shortest time, which means that the efficiency of transmitting data files is the greatest. The reason why multi-file and multi-thread

Number of threads	2	3	4	5	6	7	8	9	10
Experiment number	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)
1	33.04	29.48	28.18	28.33	32.17	35.18	34.54	38.35	31.01
2	33.23	30.96	28.73	32.90	33.06	39.80	37.20	28.91	42.49
3	42.72	30.77	31.22	33.67	28.75	37.29	31.09	31.04	29.12
4	39.88	26.29	27.98	27.36	31.62	35.05	31.59	43.43	35.57
5	38.93	31.20	31.57	27.01	31.24	35.17	32.22	37.62	34.51
6	32.73	31.01	29.80	29.44	37.11	32.64	51.20	35.72	39.52
7	29.21	29.35	27.29	30.59	28.91	39.77	44.88	32.18	40.67
8	33.55	32.52	37.22	33.61	32.39	38.48	32.5	40.73	53.68
9	36.30	29.22	28.72	33.73	30.48	31.79	37.86	34.27	40.68
10	35.15	28.10	29.87	27.80	36.47	35.65	33.49	400	48.05
11	31.09	30.33	26.81	34.81	41.74	38.31	32.29	36.81	34.48
12	30.76	29.98	29.83	31.72	28.46	32.15	33.25	40.83	49.30
13	37.06	29.19	32.46	36.89	32.56	33.06	34.97	38.64	43.59
14	33.03	29.50	29.05	28.07	35.02	39.11	33.06	37.14	33.17
15	30.25	29.55	27.83	30.29	31.58	33.21	42.12	36.53	29.37
Average	34.46	29.83	29.77	31.08	32.77	35.78	36.15	36.81	39.01

Table 4. Experimental data of single file transmission with different number of threads

transmission is slower is that operation system has different strategies to read single file and multi files, which makes the performance different. From the average transmission time of the experiment, we can get that the average network bandwidth utilization of single-file and multi-thread file transmission achieves 80.07% of theoretical value. Compared with general file transmission, single-file and multi-thread file transmission can improve the network bandwidth utilization.

Therefore, based on the experimental results, the paper gets the conclusion that the time of multi threads transmitting the same data file is the shortest and can achieve the best transmission efficiency.

According to the experimental results, before transmitting satellite data files, some little data files can be combined into a greater data file, and then single-file and multi-thread transmission mode is used to transmit data files. After the transmission, the data files are divided into the original data files according to the marks, which can save transmission time and has great application value when the satellite data volume is very large.

3.5 Selection of the number of threads

Single-file and multi-thread data transmission mode in the paper can reduce the file transmission time and improve the file transmission efficiency, but the best number of transmitting file threads need to be studied. So we continue to make experiments and increase the thread number of transmission file to verify the problem. The number of threads for the experiment is from 2 to 10. 9 schemes are compared, and the data volume transmitted by each scheme is 2 G. In order to simulate the real environment of transmitting satellite data from the southern receiving station to the data center in Beijing, the network delay of the experiment environment needs to be set. The average delay from the southern receiving station to the data center is 50ms, so the network delay in the experiment is set to be 50ms, which can meet the real transmission environment.

In order to ensure the reliability of experimental data, each experimental scheme is made for 15 times. The transmission time of each experiment is recorded, and the average transmission time of each scheme is computed. The achieved experimental result data is shown in Table 4.

The average transmission time of each transmission scheme in Table 4 is compared, and the comparison figure of transmission time is drawn, as shown in Figure 2.

We can see from Figure 2 that the average time of using 3 threads and 4 threads to transmit data files is very short, which is less than 30s. And there is no difference between them. When the number of threads transmitting data files increase, the transmission time increases, which means that transmission efficiency reduces. From Figure 2, we can get the conclusion that under the network environment, it is not real that the more the number of threads transmitting data files, the shorter the transmission time, and the suitable number of threads exist for transmission environment, which needs to be determined by making experiments.

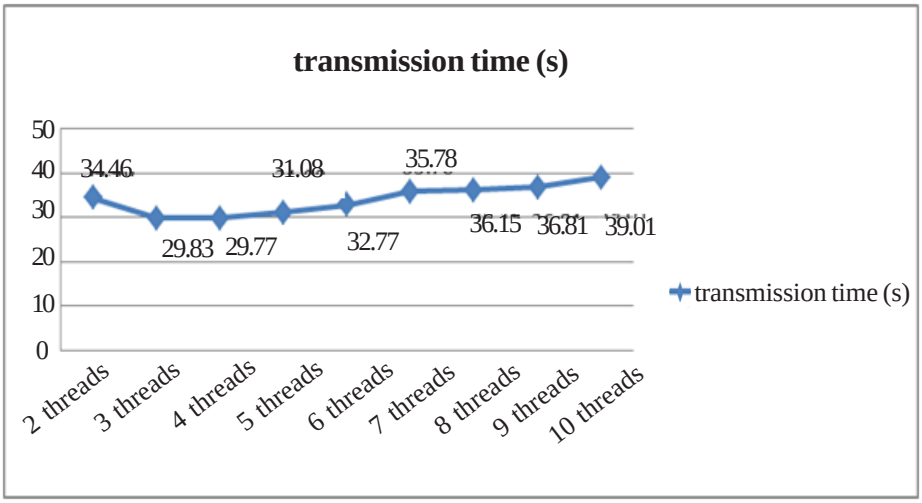

Figure 2. Comparison of transmission time of single file with different number of threads

4. Conclusion

The paper firstly makes comparison on transmission efficiency from the perspective of the number of files and threads, and makes lots of experiments. The experimental data proves that the efficiency of single-file and multi-thread data file transmission is greater than that of the other modes. The paper analyzes the selection of the number of threads transmitting single file, and 2-10 threads are used to transmit specified files. Lastly, the paper proves that under the network environment, lots of experiments need to be made to determine the optimal number of threads.

Proving the relationship of the number of threads in real network environment and transmission data file needs lots of experiments, so the problem needs to be deeply researched.

References

[1] Feng, Wang., Jiarong, Luo. (2004). Research and application of multi-thread Socket communication under Linux, *Computer Engineering and Applications,* 40 (16) 322-327.

[2] Peixian, Wu. (2005). Analysis on Socket program based on TCP under Linux, *Modern Electronics Technique,* 28 (16) 53-55.

[3] Bin, Zhang., Sun Yufang. (1997). Design of multi-thread parallelization of file server based on microkernel, *Chinese Journal of Computers*, 20 (8) 737-743.

[4] Hao, Feng., Lei, Su. (2008). Multi-thread network file transmission system based on TCP, *Software Guide*,7 (8) 93-94.

[5] Yuanyang, Wang., Yuanping, Zhou., Huanli, Guo. (2009). Implementation of multi-thread concurrent communication based on Socket under Linux, *Microcomputer Information,* 25 (5-3) 106-109.

[6] Stevens, Richard W., Rago, Stephen A. (2006). Advanced Programming in the UNIX Environment (2nd), Post & Telecom Press, p. 200-300.