§5. High-speed Data Transfer Experiments with Long-fat Network

Yamamoto, T., Nakanishi, H., Nagayama, Y., Hiraki, K., Inaba, M., Tanida, N. (Univ. Tokyo)

ITER is a next-generation experimental tokamak reactor and it is a key for realization of the fusion. After the negotiation in 2005, it was decided that ITER reactor will be constructed in France and the International Fusion Energy Research Centre (IFERC) will be built in Japan within a framework of ITER Broad Approach (BA). The experimental data which is produced by ITER experiments and the video information at the control room would be sent to the ITER parties.

In the case of Japan, SNET is a good scheme to distribute the massive data of ITER experiment to the related universities and institutes. The collaborator of the remote site could refer the data of ITER with the same method on LHD or QUEST. However, long fat-pipe network (LFN) problem will occur because the distance between Japan and France is so long[1]. Even the long network has the bandwidth of 1 Gbps, the real speed of the file transport does not reach to 1 Gbps without the adequate adjustment. TCP protocol requests the sender waits the acknowledgment of the received data for error recovery, such as packet loss. The round trip time (RTT) of the long network is so large, the waiting time of the sender is increased and then the effective bandwidth is decreased.

The WAN which covers the nation is enough long to be effected by LFN problem. For example, the file transfer is very slow (under 100 Mbps) between NIFS and Kyushu University, which its RTT is about 30 ms. To overcome this situation, there are various researches about TCP mechanism. The adjustment of inter packet gap (IPG) is the most effective method to avoid the packet loss on LFN. The precious study of the data transfer with 10 GE line are discussed in Ref. 2).

Investigation which is how to maintain the throughput high by the several real-network testing is very important. To reveal the LFN problem, the data transfer experiments between Japan and France had performed two times in 2009.

The first experiment was performed on July 2009. The servers which located at both ends were connected to 1 Gbps line. RTT between Toki in Japan and Cadarache in France is about 320 ms. The path of the experiment was the shared line, SINET3, GEANT2, and Renater, most of the bandwidth of the path was 10 Gbps. The path was quite clean, no packet loss was observed through the experiment. No special hardware was required to build the server; the server at Cadarache was standard notebook PC which has 1 GE network interface card (NIC). We had transferred more than 1 TB of data in three hours from memory to memory with the technique of avoiding the packet loss, which was the adjustment of IPG by the modified NIC module driver. The average throughput was 880 Mbps.

On September 2009, the second experiment has performed. In this experiment, we used 10 Gbps line and the target was not only memory-to-memory transportation but also disk-to-disk one. The servers were connected to both the shared line (SINET3 path) and the dedicated line (JGN2plus/APAN-JP path) in Fig. 1. The bottleneck of 4 Gbps existed in Europe on the both path. We could transfer 86 GB of data in 205 seconds between the disk at Toki and the disk at Cadarache through both the shared line and the dedicated line. The average throughput was 3.3 Gbps, which was more than 80 % of the limitation. These experimental results encourage us to prepare the next step of the virtual laboratory.

Fig. 1 The network path between Toki and Cadarache.

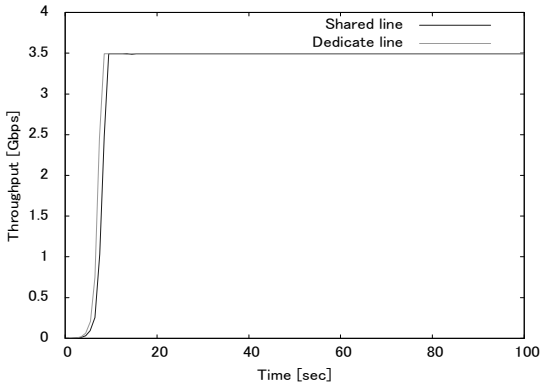

Fig. 2 The throughput of 2nd experiment. The available bandwidth was limited to 4 Gbps.

1) T. Yamamoto: Fusion Eng. Des. **83** (2008) 516-519.
2) T. Yoshino, et al.: Proc. of the 2008 ACM/IEEE conf. on Supercomputing (2008).