Design and performance of the network system for the storage ring control at SPring-8

T. Fukui, K. Kobayashi, T. Masuda, A. Taketani, R. Tanaka, T. Wada and A. Yamashita
SPring-8, Kamigori, Ako-gun, Hyogo 678-12, Japan

Abstract

This paper presents the design and performance of the network system of the SPring-8 storage ring control. The network system consists of five Ethernet/FDDI switching hubs and a router interconnected by dual loop FDDI. VME CPU boards and operator consoles distributed over the accelerator complex are connected to the switching hubs through optical fibers or metal Ethernet cables. The LAN switching technology is applied to the system, and it provides good performance for the accelerator operation. The remote communication between the VME CPU boards and consoles are performed by the remote procedure calls(ONC/RPC) with TCP/IP protocol. The data transfer rate between VMEs and consoles by the RPC is around 300~500kB/sec. The average latency including overhead of the control software is measured to be 6msec which is good enough for the storage ring control.

1 Network system design

Figure 1 shows the SPring-8 storage ring network system. The system consists of five Ethernet/FDDI switching hubs and one router. Each hub or router is connected by a dual loop FDDI as the backbone. We choose the FDDI because of its robustness when using dual homing connection and the wide bandwidth (100Mbits/sec). If one connection fails, it automatically enables the link to another connection. The router is used for the gateway to the injectors(Linac and Synchrotron). The switching hubs have two functions. One is a bridge between FDDI and Ethernet and the other is a pier to pier connection between the consoles (HP-UX) and the VME CPU boards (HP-RT, HP version of LynxOS) with shared backbone. This configuration allows for each console or VME CPU board to occupy the full bandwidth of the Ethernet (10Mbits/sec) with negligibl few Ethernet packet collisions. We use HP9000 C100 and 712/60 for the consoles, HP9000 743rt for VME CPU boards. The connections between the switching hubs and the VME CPU boards use optical fiber in order to avoid electro-magnetic interference.

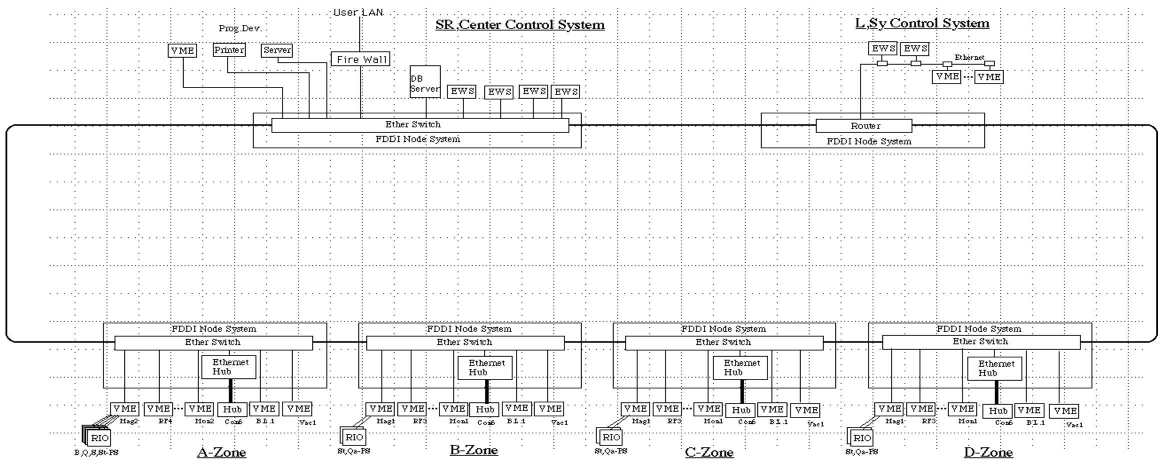

Figure 1. The network system for the SPring-8 storage ring control system.

The connections between the switching hubs and the consoles use unshielded twisted-pair cables except for the database server. The database server is connected by the single loop FDDI to keep wide bandwidth. For maintenance of the VMEs and the equipment in the storage ring, the 19 normal hubs are distributed around the storage ring and they provide more than 100 Ethernet ports around the maintenance corridor and the storage ring tunnel. The storage ring control network is isolated from the laboratory public LAN by the firewall and we use the private IP address[1]. The network traffic is controlled as to allow the program development and the replicated database access[2].

2 Remote communication

We use the remote procedure call (RPC) to communicate between the consoles and the VME CPU boards. The RPC is chosen for two reasons. One is the ease in developing networked applications with the distributed computing environment. The other is that it makes the client/server model of computing much easier to program.

There are two important commercial implementations; Sun’s Open Network Computing (ONC) and OSFs Distributed Computing Environment (DCE) based on the HP/Apollo NCS RPC. Currently, the two implementations are incompatible. We choose the ONC/RPC because it is widely available for many platforms including the real-time OS, HP-RT. For the network transport protocol, we choose TCP/IP. In general, UDP has better performance than TCP and TCP is more reliable than UDP. The reliability of TCP means that if a reply message is received, the procedure is executed exactly once. The performance advantage of the UDP is small and the reliability is more important for our control system.

3 The performance of the network system with the RPC

Figure 2 shows the data transfer rate of the RPC with UDP and TCP between the console (HP9000 C100) and VME CPU board. The difference between the UDP and the TCP is less than 10% up to 4kbytes of data. We found a large delay with data sizes around 5kbytes. It comes from the standalone ACK timer of TCP’s implementation (200msec). This timer works to keep the minimum number of packets to send. For example, a telnet session sends a character to the TCP protocol layer at each key typed, if the layer sends a character

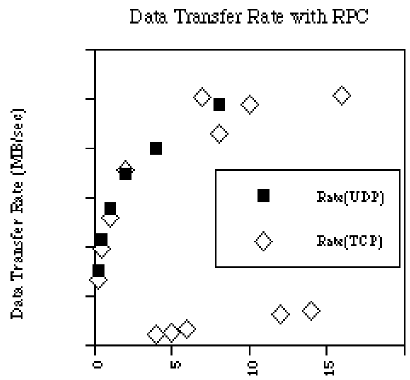

Figure 2. The data transfer rate between the console and the VME CPU board with the ONC/RPC.

to the Ethernet layer, the network traffic becomes very busy The TCP protocol layer waits for a character during an interval of the standalone ACK timer. Most of the time, the small size of the data is not affected from the standalone ACK timer but the large delay happens at a rate of less than 10^{-4}. In this case, maximum delay is 200msec for one way and it is possible to delay 400msec for the RPC at the rate of 10^{-8}. For the large size of the data, the data transfer rate becomes more than 500kbytes/sec and this is about half of the full bandwidth of the Ethernet.

4 The performance of remote communication of the SPring-8 storage ring control system

The framework of the network communication for the SPring-8 storage ring control software sends 256 bytes data to the Equipment Manager (EM) [3] from the Access Server(AS) on the console. The EM is the RPC server and runs on the VME CPU board, and the AS is the RPC client. The AS receives the message from the Graphic User Interface via the Message Server[4]. The message contains the equipment name and the AS resolves the server name to access the equipment from the database. The AS can handle multiple servers and sends the message to multiple servers simultaneously. The multiple servers run synchronously. Figure 3 shows the typical data transfer time including the overhead of the AS and the EM, and the average is 5.8msec. The EM takes less than 0.1msec because of using the dummy function and the AS takes about 2msec and the RPC call is around 4msec. The network communication is in good agreement with the data of the Figure 2.

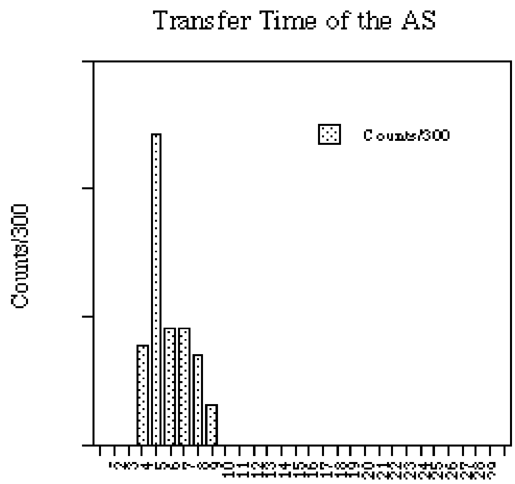

Figure 3. Transfer time between the AS and the EM. 256bytes data is sent to the EM.

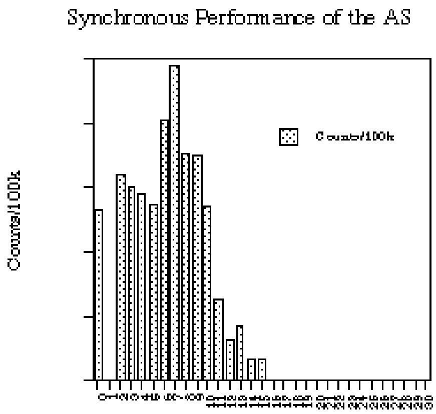

Figure 4. The time difference between the first received EM and the last received EM. 256bytes data is sent to the 4EMs from the one AS simultaneously.

Figure 4 shows the synchronous performance for the AS to send 256bytes data to the 4 EMs. The time is defined as the difference between the first data received by the EM and the last data received by the EM by sending from the console at the same time. The average time is 7.8msec and only 1 event takes more than 30msec. This is good enough for the storage ring control system. For example, if we change the current of the 576 steering magnet power supplies around the storage ring, the setting of all power supplies is finished within 8msec most of the time.

5. Summary

We designed the network system of the SPring-8 storagering control. The LAN switching technology is applied to the system, and it provides good performance for the accelerator operation. The data transfer rate between the VME CPU boards and the consoles by the ONC/RPC is around 300~500kB/sec. The average latency and synchronous performance including overhead of the AS is measured to be 5.8msec and 7.8msec which is good enough for the storage ring control.

Reference

[1] The Internet Assigned Numbers Authority, RFC1597.

[2] A.Yamashita et al., "The database system for the Spring-8 storage ring control", in these Proceedings, ICALEPCS'97, Beijing, China, 1997.

[3] A.Taketani et al., Proc. of ICALEPCS'95, Chicago, USA, 1995, p.625.

[4] R.Tanaka et al., Proc. of ICALEPCS'95, Chicago, USA, 1995, p.201.