MASS STORAGE SYSTEM BY USING BROADCAST TECHNOLOGY

HIROFUMI FUJII, RYOSUKE ITOH, ATSUSHI MANABE, AKIYA MIYAMOTO, YOUHEI MORITA, TADAO NOZAKI, TAKASHI SASAKI, YOSHIYUKO WATASE

National Laboratory for High Energy Physics (KEK), 1-1 Oho, Tsukuba, Ibaraki 305, Japan

TOKUYUKI YAMASAKI

Broadcast Products Company, Sony Corporation, 4-14-1 Asahi, A tsugi, Kanagawa 243, Japan

There are many similarities between data recording systems for high energy physics and broadcast systems; e.g., 1) the data flow is almost one-way, 2) requires real-time recording, 3) requires large-scale automated libraries for 24-hours operation, etc. In addition to these functional similarities, the required data-transfer and data-recording speeds are also close to those for near future experiments. For these reasons, we have collaborated with SONY Braoadcast Company to study the usability of broadcast devices for our data storage system. Our new data storage system consists of high-speed data recorders and tape-robots which are originally based on the digital video-tape recorder and the tape-robot for broadcast systems. We are also studying the possibility to use these technologies for the online data-recording system for B-physics experiment at KEK.

1 Introduction

In high-energy physics experiments in near future, high-speed and large data storage systems are required. For example, B-physics experiment at KEK requires recording speed of 15 MBytes/see and the capacity of 120TBytes storage. For these requirements, we have interested in broadcast devices, e.g., digital image recorders, video tape libraries, etc. These devices have possible capabilities for high-speed and large-volume computer storage devices although these are maily designed to record and replay video images. For this reason, we have started the collaboration with Sony Coporation to apply these devices as computer storage devices. The preliminary studies have been started since 1992 [1,2] and official collaboration has been started since 1994.

2 Broadcast devices

The broadcast devices we have interested in are the followings;

- digital image recorders
- video tape libraries
- serial digital data transfer devices

these devices are maily used for commercial video advertizements; many short movies are recorded on digital video tapes, and these tapes are stored in the tape library. Each movie is automatically replayed according to the broadcast schedule. This situation is similar to the high-energy experiments. The event data are recorded on tapes, and these tapes are stored in the tape library. Each tape is automatically mounted to the tape device according to the requests from the processes. Of course, there are many differences between advertizement movies and our data. In the following sections, the differences and the modifications for our purpose are discussed.

3 Digital Image Recorder

A digital image recorder is essentially a video recorder. To use this device as a computer tape device, the following problems had to be solved.

- Standard computer interface is needed. The device has parallel input/output data ports and synchronous data-flow is requested. On the other hand, in general, computer read/write operations for external devices are asynchronous.
- Because that the high-speed image recorder is a helical-scan drive, it takes long time for restarting the servo-motor and for re-positioning the tape.
- The tape-device has no structure; no end-of-file mark, no inter-record gap. However, digital data tape has an ID track for high-speed positioning.

Items (1) and (2) can be solved by preparing enough buffer-memory in the interface. Actually, there are several 3rd vendor products of computer interfaces for DIR-1OOO which has large buffer-memory.

The Sony SCSI interface has also have large buffer memory. The interface is designed as SCSI-to-ECL parallel interface so that no change is necessary on the drive. This interface also has lookup table for high-speed positioning. The table is also written at the beginning of the tape and loaded into interface memory at tape-loading time. If the table is changed, the updated table is also written when unload the tape. This also allows us to have a file structure on the tape.

The Sony SCSI interface can be controlled through the same SCSI port for data 1/0. This is convenient for small stand-alone systems. However, for large scale storage system, it is necessary to separate the control path from data path. For this purpose, the RS-232-C port can be used as an alternate control path.

4 Tape Library

The tape library for broadcast system had the following problems for our purpose;

- The number of drives in the library is quite limited (e.g., 3 drives).
- The capacity cannot be expanded.

The new tape library has been designed according to the following guidelines;

- The number of drives can be increased.
- The capacity can be increased
- Any tape in the library can be mounted at any drive.

The new tape library is designed based on the DMS-24 with additional direction of the robot. The new tape library also has junctions to connect shelves and through these junctions, tapes can be transferred between shelves.

5 Data Transfer Interface and Switches

For high-speed digital data transfer, a serial digital interface is used in broadcast systems. The data format is standardized in SMPTE-259M. The transfer speed is 270 Mbps. The routing switchers are commercially available; e.g., Sony DVS-V series can switch up to 64 x 64 as in single unit, and up to 512 x 512 when cascaded. The advantage of this device is that the interface is designed to transfer the digital video signals so that it is easy to use for digital video recorders. However, at present, the transfer speed is limited to around 270 Mbps and it is not designed to use as a computer device. For these reasons, we have chosen computer network devices for the data transfer.

6 Softwares

6.1 *SONY File Bank System*

The SONY file bank system is an NFS-based data management system. The control workstation (NFS server) is used for an interface between NFS and the tape library system.

6.2 *OSM*

The Open Storage Manager (OSM) is tested at Sony Technology Center. The program has successfully customized for Sony tape drives and robots.

6.3 *CERN-SHIFT*

CERN-SHIFT program is tested at KEK. The tape server part has been modified for Sony tape drives and robots. The remote tape copy and disk pool manager successfully run on the Sony tape system.

7 System Examples

7.1 Storage System for KEK Central Computer

The new central computer system at KEK has two tape libraries, the one has 12 drives in the library and the other has 8 drives. The capacities are about 7.5 TB and 12.5 TB respectively. The tape drive is based on a digital video recorder and maximum recording speed is 12 MB/see. Because that this system is used for rather

small groups, there is no need for the tape transfer between libraries. The system is developed by Hitachi and Sony collaboration. This system will be in operation from Jan. 1996.

7.2 Data Recording System for *KEK* B-experiment (BELLE)

Because the required maximum recording speed for KEK B-physics experiment (BELLE) is about 15 MB/see, Sony DIR-1OOOM and its SCSI-2 Fast/Wide interface is one of the candidate as a data recorder drive. In this case, the requested capacity is more than 30TB/year and library-to-library tape transfer is requested for long term operation. Sony proposed us a couple of methods which is easy to build by using current technology.

8 Summary

We have collaborated with Sony Corporation to study the possibility to build a high-speed and large-scale mass storage system by using broadcast devices. The system according to these studies will be in operation soon at KEK.

References

1. H. Fujii et. al., *An ID-1 Mass Storage Project at KEK*, Proceedings on CHEP'92, pp.711-714.
2. Y. Morita et. al., *ID-1 Mass Storage System for Mainframe by using FDDI network,* Proceedings on CHEP'94, pp.219-221.