§22. Object-oriented Data Operation and Storage of the LHD Mass Data Acquisition System

Nakanishi, H., Kojima, M., Ohsuna, M., Komada, S., Emoto, M. and *LABCOM Group*,
Sugisaki, H. (CTCS)

The LHD data acquisition and management system for plasma diagnostics[1, 2], which is named as *LABCOM system*, has been designed to process ~ 1 GB plasma raw data by the typical discharge experiment of about 10 second short-pulse duration in the continuous 3-minute cyclic operation. In order to satisfy the necessary conditions to complete to acquire and process whole diagnostic data within 100 seconds after every discharge end, it was obliged to utilize the massively parallel processing (MPP) structure and reduce the data processing load of the individual element, such as data I/O ports or cpus. After about a year LHD experiment, the total amount of the raw data from kinds of the plasma diagnostics goes up to about 250 MB per each shot, than 20 ~ 30 MB in the start-up experimental period.

CAMAC Data Acquisition on Windows NT

The primary purpose of the data acquisition part of the LABCOM system is the plasma measurements of the physical analysis which generally requires as fast data sampling as to follow the plasma intrinsic frequencies of its dynamic behaviors. They manages the CAMAC digitizers and transfer, transform and store the experimental data, where all of the processing sequences are executed in the pre- and post-experimental periods as they were pre-programmed. These procedure flow of the data management are often recognized as so-called the batch processing.
Sub-components of the LABCOM data acquisition can be classified into the categories as listed below,

1. digitizing by CAMAC modules
2. SCSI data transfer with optical extender
3. PC/Windows NT data acquisition server computers
4. data cross-referencing by virtual shared memory
5. object-oriented database on every server computer
6. harddisk array and media-changeable mass-storage

As for the digitizer modules for LHD, we have to inherit a lot of CAMAC properties which had been working in the prior experimental devices of our institute.

In order to manipulate the SCSI-connected CAMAC crate controller from Windows NT and to acquire the data from their modules, we have newly developed the CAMAC handling software[3]. The block data transfer between the CAMAC governing service program and the client application will be executed efficiently through the double buffering mechanism. The total throughput of the block data transfer is about 700kB/s between the CAMAC module and the application program.

Hierarchical Data Storage

The LABCOM system has applied the hierarchical storage management (HSM) technology. It had installed the 2 levels of the storage hierarchy in 1998: One is the local RAID of 50 GB for each server computer for 30 kinds of diagnostics, and the other is the 3 MO jukebox whose total volume goes up to 3.6 TB.

The 50 GB RAID is occupied by the database volume file of the object-oriented database management system (ODBMS), and the plasma data are directly store into the database volume as the data object. The typical data storing rate into the ODBMS is max. 400 kB/s, and to improve the effective rate and reduce the storage volume size, the data compression mechanism of *GNU-zlib* library is used toward those binary mass objects. Owing to the data compression and multi-thread programming, almost all of the server PCs can finish their acquisition and storage within 40 seconds after every experiment end. The compression rates, however, tightly depend on the signal forms; some integrated signal or pulse-count measurements achieve fine rates about a few %, while some fluctuation signals never go under 50 %.

Operational Results

The LABCOM system has been successfully developed with both of the strong capability of the mass data acquisition and the expansion flexibility by applying the MPP structure based on PC and Windows NT. At the end of the 2nd experimental term, the total amount of the plasma raw data acquired by the LABCOM system was about 120 MB/shot, excepting some kinds of the standalone acquisition systems. In the following 1 or 2 years, it will easily reach the initial estimate of ~ 1 GB/shot.

Table 1: Operational achievement of the LABCOM system in the 1st and 2nd LHD experimental terms. 1st:#1~#1888, 2nd:#1889~#7132.

	1st term /MB	2nd term /MB
Halpha	4,231	41,674
Bolometer	8,276	51,114
Fast Ion Gauge	1,825	7,273
Impurity Monitor	2,169	21,568
Magnetics	12,847	90,300
Langmuir Probe	6,706	58,050
mm-wave Interfero.	4,361	-
ECH Power	3,380	-
Reflectmetry	-	23,587
Soft-X fluc.	-	25,314
X-ray PHA	-	9,420
ICH Power	-	15,577
Bremsstrahlung	-	28,162
Fast Ion	-	6,860
TOTAL	43,795	378,899

References

[1] Nakanishi H. *et al.*, J. Plasma Fusion Res. **72** (1996) 1362.
[2] Nakanishi H. *et al.*, Fusion Eng. Design **43** (1999) 293.
[3] Kojima M. *et al.*, Fusion Eng. Design **43** (1999) 433.