

§30. "LABCOM/X" Experiment Data Platform and Fusion Virtual Laboratory

Nakanishi, H., Ohsuna, M., Kojima, M., Nonomura, M., Imazu, S. (Pretech Corp.), *LABCOM Group*

The data acquisition and management system of LHD has been operating for the past 12 years. At the beginning of LHD experiment, it started with only several CAMAC diagnostics whose number was less than 10. However, it soon went up to 30 even though all of them were still CAMAC digitizers, and after we made the long-pulse trials in earnest, the real-time data acquisitions have much increased to reach 75 nowadays (Fig. 1).

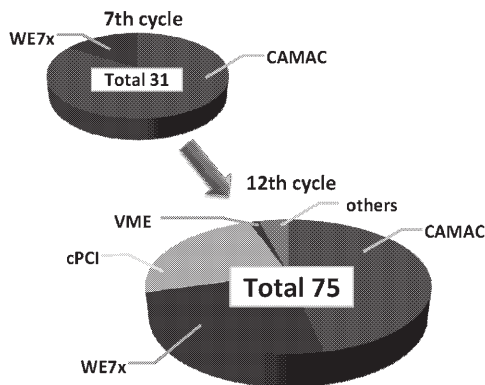


Fig. 1. Digitizer variation in LHD:

All the CAMAC data acquisitions are still working in LHD, however, more than half of the total DAQs are using real-time capable digitizers, such as Yokogawa WE7000 and National Instruments PXI or namely CompactPCI. The acquired data amount in a short-pulse experiment grows up to 7 Gbyte/shot presently, and the maximum record is 90 GB per shot in about 1-hour long-pulse experiment established in 2006. LHD constantly have about 170 shots every day (Fig. 2). LHD's 12-year operation has made the data amount a hundred times bigger than its beginning.

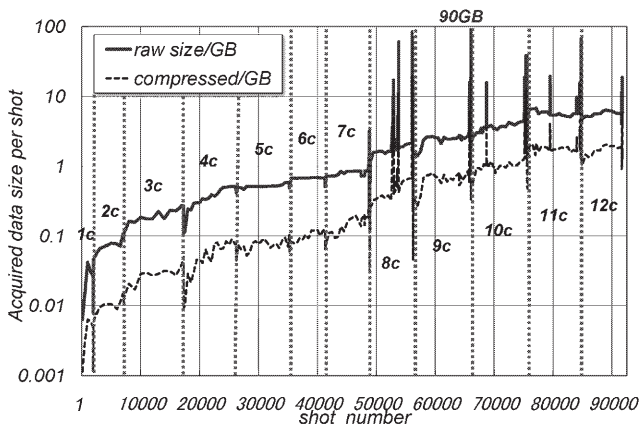


Fig. 2. Growth of acquired raw data amount per shot in LHD

Fusion Virtual Laboratory in Japan

In addition to LHD, the LABCOM system has recently covered QUEST and GAMMA10 experiments under the new framework of "Fusion Virtual Laboratory" where users can access the data equivalently regardless of their whereabouts. Such the activity is named "SNET", which is based on a closed VPN on Japanese academic internet backbone SINET3 and covers multiple experimental remote devices, at first LHD in NIFS, and QUEST in Kyushu University, and GAMMA10 in University of Tsukuba. SNET also covers remote use of supercomputer. Details are explained by Yamamoto in today's poster session. (Fig. 3)

For the experiment data platform, the functionality of data access control is mandatory when used across multiple experimental sites. It should be for both the user and data belongings for each site.

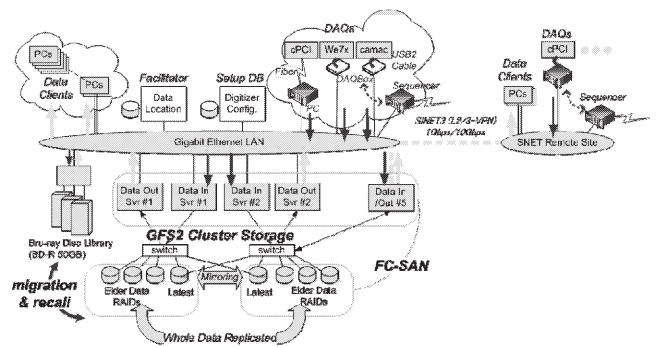


Fig. 3. Schematic view of "LABCOM/X" distributed data platform and its extension for remote site by SNET

There are 3 query keys to retrieve data in our system; site group, diagnostic or data name, and the shot number. Only the site group name must be unique in use of multiple experimental sites. In our implementation, the access privilege is given to the prior registered host addresses for each site. Of course, a collaborators can joins the multiple site groups and register his/her host to both sites to access the both data. (Table 1)

We have to handle growing enormous amount of data forever, so the light-weight processing should be the key concept for the sustainable data platform.

Table 1. Data retrieval/query keys

| site | diag. | shot # | user |
|---------|-------|--------|-------|
| LHD | diag1 | 1 | addr1 |
| | diag2 | 2 | addr2 |
| | ... | 3 | addr3 |
| QUEST | diag3 | 1 | addr3 |
| | diag4 | 2 | addr4 |
| | ... | 3 | addr5 |
| GAMMA10 | diag2 | 201234 | addr2 |
| | diag5 | 201235 | addr3 |
| | ... | ... | addr6 |

1) Nakanishi, H. *et al.*: Fusion Eng. Des. **83** (2008) 397.