

§23. Development of Index Server System for Analyzed Data Files

Ohdachi, S., Watanabe, K., Emoto, M.

Advances in diagnoses are key elements for the better understanding of the fusion plasma. No one can discuss the transport phenomena quantitatively without detailed measurements of profiles, for example. Need for the finer spatial and temporal resolution in measurements has increased the total amount of data to be acquired and to be stored. Since hundreds of mega bytes data should be handled in a typical LHD discharge, it is not realistic to do with a single computer. Fully distributed computers (PC with windows NT) were adopted for the data acquisition system for LHD under these circumstances; data are now acquired from the CAMAC crate and stored in parallel within less than 3 minutes in the system.

On the other hand, it has been difficult to access data on a different computer in this system. Since we do not share a central computer for data analysis, each data has been analyzed on distributed computers with different languages and different operating systems; it is not easy to make shared libraries to access the commonly used data. Our project – to develop an index server for analyzed data – is to make it easy to access analyzed data which is widely needed in experiments. We do not treat raw data. Our standard system (LABCOM system) has provided us a sophisticated way to access raw data on other computers.

We use text-based files as a media to distribute the analyzed data. What we prepare for the smooth handling for the analyzed data files are:

- 1) Unification of the file format[†]. It contains information about the file itself; shot number, name of measurement, number of measured points, etc. With the help of these information, analyzed data in the files can be accessed easily using standard subroutines.
- 2) To provide basic tools to organize the unified files. A database server on network manages the location of each analyzed files stored in distributed computers on the network. The name of the measurement and the shot number are used as keys in the database. Clients for the database can do the following actions[‡]: “to register a data file”, “to cancel a registration”, “to search for a file with conditions” and “to make a local copy for a certain file”.

The flow of the data in the experiments is summarized in Fig.1. Raw data are acquired by distributed systems and analyzed on different computers (A) in Fig. 1.). Those who want to provide their analyzed data, store their data in unified format and register its location to the index server (B).

The registered file can be used by anyone who wants to use that data. One can get the location of the file from the index server and can make a local copy of the file (C). Analyzed data contained in the files can be utilized with prepared subroutines.

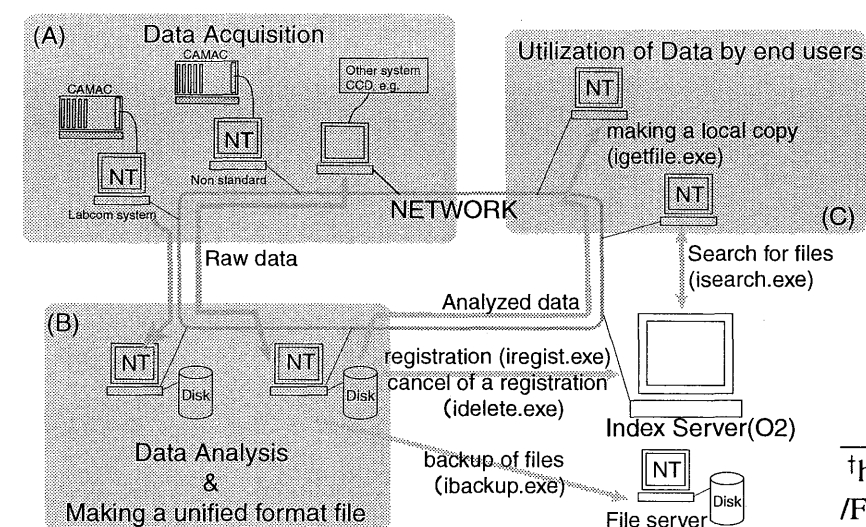


Fig. 1: The concept of this system.

Basic data, e.g. the electron temperature profile, had been registered for test of the system. Refinements of the system for the third experimental campaign are in progress.

[†]http://dgeg3.nifs.ac.jp/Software/kaiseki/File_Format_Specification.doc

[‡]Url for the information about the client programs: <http://dgeg3.nifs.ac.jp/Software/kaiseki/index.html>