Upgraded data service system for HT-7 Tokamak

QU Lian-Zheng[1,2], LUO Jia-Rong[1,3], WEI Pei-Jie[1], LI Gui-Ming[1], CHENG Ting[1], QI Na[1]

([1] Institute of Plasma Physics, the Chinese Academy of Sciences, Hefei 230031; [2] Institute of Communication Command, PLA, Wuhan 430010; [3] Dong Hua University, Shanghai 200051)

Abstract A data service system plays an indispensable role in HT-7 Tokamak experiment. Since the former system doesn't provide the function of timely data procession and analysis, and all client software are based on Windows, it can't fulfill virtual fusion laboratory for remote researchers. Therefore, a new system which is simplified by three kinds of data servers and one data analysis and visualization software tool has been developed. The data servers include a data acquisition server based on file system, an MDSplus server used as the central repository for analysis data, and a web server. Users who prefer the convenience of application that can be run in a Web Browser can easily access the experiment data without knowing X-Windows. In order to adjust instruments to control experiment the operators need to plot data duly as soon as they are gathered. To satisfy their requirement, an upgraded data analysis and visualization software GT-7 is developed. It not only makes 2D data visualization more efficient, but also it can be capable of processing, analyzing and displaying interactive 2D and 3D graph of raw, analyzed data by the format of ASCII, LZO and MDSplus.

Keywords Tokamak, HT-7, Data system, MDSplus, GT-7

CLC numbers TL375.1, TP302.1

1 Introduction

HT-7 is a superconducting Tokamak in China used for controlled nuclear fusion research as a national project.[1] There are some shortcomings in the former data system for HT-7. Firstly, it is one file system based on binary format. It cannot compress the acquisition data and it is difficult to share signal configuration and analysis result. Secondly, it provides no interface to commercial software such as IDL, Matlab and Labview. Researchers have to use a simple visualization software tool based on Microsoft Windows to explore signal data. While processing and analyzing the raw data, they have to use the upper software to transfer the binary data to text format; also they can't add any comments to server. Thirdly, HT-7 has a heterogeneous group of users. Some of them prefer Windows, some Linux, some UNIX and some Macintoshes. Some researchers can program in a high-level computer language and others not. Some scientists on HT-7 are full time and others not. But the former data system only provided softwares based on Windows. Obviously, it is very inconvenient for the researchers who are not familiar with Windows.

In order to resolve these problems, the authors have developed a new data service system as explained in part **2**. Researchers can get the data with a unified data retrieval method no matter what operating system they are using and where they are.

2 Data system overview

Fig.1 shows an overview of the HT-7 data system. It is divided into two parts. One part is DAS (data acquisition system) and the other is DSS (data service system). The DAS is a batch system. It consists of more than 20 Microsoft Windows based industrial control computers. After each pulse, the raw data which come from digitizers will be compressed by the format of LZO[2] and saved to local disk, then the configuration of signal such as calibration of the instruments, accuracy and speed of data acquisition delivers

Supported by National Natural Science Foundation of China (No.19981303)

Received date: 2005-01-25

will be written into a defined data structure and attached to the data file, and then the LZO file will be sent to the main server by TCP/IP protocol. When all data of a pulse have been transferred to server, program will check disk space and delete the old data according to relational algorithm. At the same time all data will be directly sent to analysis server to be processed at a low level. When receiving a new shot number from shot number server, program running on the MDSplus[3-9] server will start threads to do many kinds of calculation and processing. Each thread implements a kind of calculation, and then writes the result into MDSplus tree automatically. This program is developed based on GNU C++. At the same time there is another calculation program running on web server. While receiving a new shot number, it begins to calculate some eigenvalues such as length and height of plasma current, plasma density and then saves the result into MySQL database. Some eigenvalues need to be input by hand.

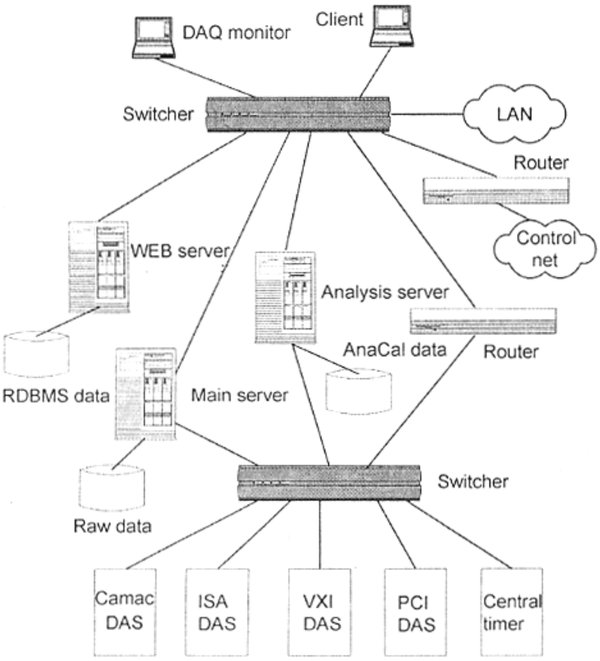


Fig.1 Data system overview.

2.1 Main server (acquisition server)

All the data are organized by shot number and signal name on main server. During experiment, a directory named by shot number were created after pulse and all the data named by signal name were saved under the directory.

Main server has taken the first place in data system since 1995 when HT-7 began to discharge for the first time, and it will be available in the future for the following reasons. Firstly, in order to save the disk space, all the data are compressed by algorithm of LZO, which provides fast speed of compression and decompression and satisfying compression rate. Performance of compression will be shown in performance test section. LZO not only supports many operation systems such as Windows, Linux and Macintosh, but also provides many libraries for different programming environment such as C++, Pascal and Java.[2] Only when knowing what version of algorithm is used, can researchers easily operate the data. Secondly, the present data acquisition system is integrated and stable. Data can be transferred to server in time after pulse. Thirdly, many programs have been designed to get this format of data, if not, they can't work. Lastly, present data system is easily extended for long pulse experiment especially for EAST, which is being constructed by separating a large data file into several small ones.

2.2 Analysis server

MDSplus is used widely in fusion labs such as CMOD, DIIID, JET and NSTX.[5] It provides enough libraries to maintain data and interface for commercial software such as IDL, Matlab and Labview. Users can easily retrieve data from MDSplus server with simple routines from MDSIP service. Therefore, the authors have adopted MDSplus as data analysis database since the spring experiment of 2004. Fig.2 shows the MDSplus tree structure for HT-7.

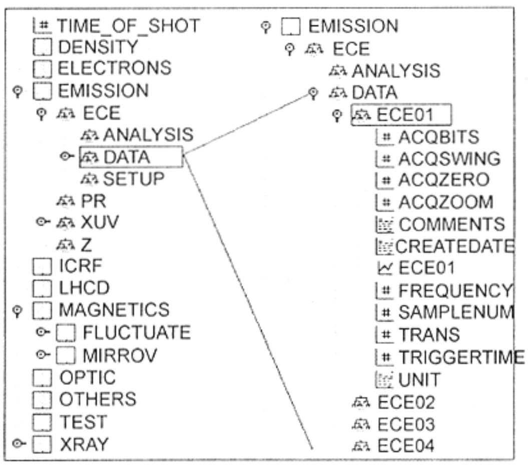


Fig.2 MDSplus tree structure.

All existing data are written into MDSplus trees. Each tree has the same structure. Below the top node there are ten subtree nodes, which are classified by the type of signal. Under the nodes, many structure nodes are built for the same kind of data such as electronic density signal. Each upper node consists of three nodes which are data, setup and analysis. The data node is used to store raw data, while the setup node is used to save configuration and the analysis node is used to store expression or analysis results.

Access control to these trees was implemented by the standard MDSplus method based on mapping remote usernames and IP addresses to local usernames. The public database does not require any writing access reading restrictions. Read access to working database was enabled for registered users. Specific users at each site were given relational permission to write data directly.[7]

2.3 Web server

IIS based on Windows was chosen to act as web server. The web site has direct HTTP access to the data directories and allows the user to read it. The website contains two parts, one is MySQL database which provides shot, information and experimental status in detail such as length of IP, method of heating, the power of LHCD and status of shot, and the other, which is a java applet, provides standard methods for viewing and manipulating data on acquisition server and analysis server. When a registered user accesses experimental information, data will be got from database automatically and displayed by format of table. Also the server allows users to query all the information they want by SQL. When users access acquisition data, they only need to input shot number and signal name or multi-shot number and multi-signal name, and then the java applet will be downloaded to client computers and get data from the upper two servers by an index program which is running on the web server. The index program will automatically judge where the data exist and then transmit them to the applet to plot. In order to improve accessing speed, the authors have provided two methods. One is full data view and the other is sampling data view. User can make any choice to access each one according to net band.

3 User interface

Fig.3 shows how to access data from data system. The X-Windows based program named DwScope is supplied by MDSplus, and it is powerful and efficient for any number of single *x*-*y* traces per frame. Some users are willing to give up some of the features and performance of the MDSplus tools and IDL codes for simpler point-to-click web pages to access their data, and they can use the web server with counts and passwords. Some offline users can download the data files to local and process them anywhere. Also users can program their own codes to access and process data. The authors have designed a universal data analysis and visualization software named GT-7 for LZO data, MDSplus data and text data with VC++6.0 based on Windows.

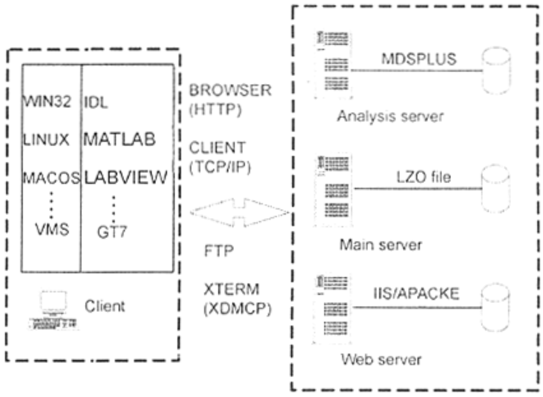


Fig.3 Modes of accessing data.

GT-7 is the main data visualization software. It is used heavily during experiment operation and in-depth analysis. It can show any *Y* versus any *X* and 3D plot for signal data array and it has the ability to plot data of different shots to compare with different colors. Fig.4 shows the main window of GT-7. Menu options are provided for most common operations: smoothing, filter, integration, differentiation and fast Fourier transforms such as calculating amplitude spectrum, phase spectrum, power spectrum, and time-frequency spectrum and so on.

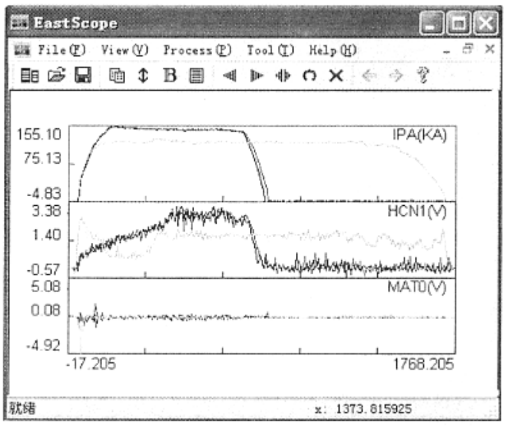
Fig.4 (a) gives a comparison of signal data among three shots. All signals share common *X* axis which stands for time sequence, and every signal has its own *Y* axis, which is physical amplitude.

Fig.4 (b) shows the organization of the displaying window for raw data, processing data and analyzing data. The full view is divided into three parts. Left view shows 2D plot for four raw signals, and the right top view shows processing result of a signal. Users can right click mouse and choose processing method such as smoothing or filter to process data deeply. The right bottom view shows the analysis result such as integration, differentiation and computation of every kind of spectrum.

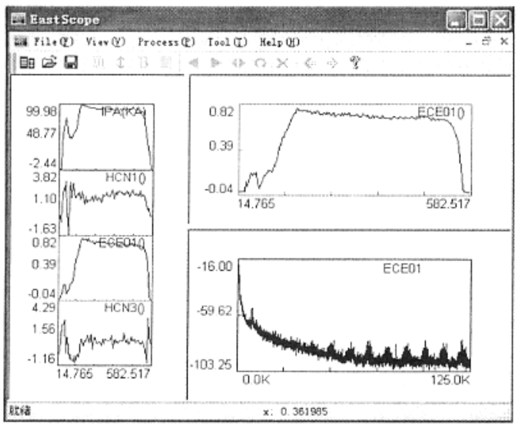
Fig.4 (c) demonstrates the organization of the displaying window for data array. The full view is also divided into three parts. Left view shows 2D plot for four raw signals at fixed position, and the right top view shows a profile slice for signal array at fixed time with different colors, and it also shows several profiles for such signals at different fixed time and different shots. The right bottom shows 3D plot for the signal array and the plot can be scaled and rotated with interactive manner, *X*, *Y* and *Z* axes respectively stand for time, space position and physical amplitude dimension.

Users can choose a plot or trace for manipulation by clicking near item of interest. Once a plot is selected, its appearance can be altered and a red rectangle around the plot can be seen, and the plot can be scaled with mouse to see the overview or detail. In all of above views, users have the choice of having crosshairs following the mouse cursor, displaying all plots on the coordinates of the cursor in the reference frame of the plot in which it resides, and users can easily see the information of a signal in detail from a popup menu choice.

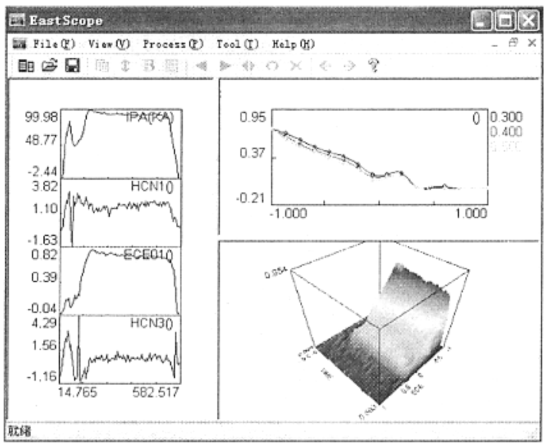
GT-7 is also able to plot any mathematical combination of data signals from acquisition server or MDSplus. For example, if users want to calculate the sum of “IPA” and “IPB”, they can input the expression of “\IPA+\IPB” instead of a signal, and then the program will analyze signals named by “\” and finish calculation immediately after retrieving data from the server. If signal data have different frequencies or length, program will automatically unite those of them at first.



(a) Signal contrast.



(b) Signal analysis.



(c) Interactive 3D vertex and 2D slice plot.

Fig.4 Main window and function of GT-7.

The analysis results with GT-7 can be saved to disk by text format and the plots can be printed, saved as BMP file or copied to the third software for being manipulated deeply. A feature especially useful during experiment operation is the ability to track the most recent pulse, so that users can create a set of plots that automatically refresh every shot. For users unfamiliar with signals available in HT-7 repository, a help menu is provided, along with the ability to link a PDF file and a web server to display more detailed information on a signal of interest. All of these features combine to make data visualization more efficient and more productive.

4 Performance test

In order to estimate the performance of the data service system, many tests have been done.

Table 1 shows performance of data compression, including MDSplus and LZO. Table 2 shows the time step sequence from the request for physics data to visualization of the whole data access system, including web and DwScope which is a data display software embedded in MDSplus and GT-7. Table 2 also shows performance contrast between the new and the old. The test data size is 13.96MB, which is the total size of 7 signals. The computer used for the test is shown in Table 3.

Table 1 Performance of data compression

Shot No.	Size (MB)	LZO			MDSplus		
		LZO size(MB)	Compress time(s)	Decompress time(s)	MDSplus size (MB)	Compress time(s)	Decompress time(s)
50014	130.5	49.2	14.1	13.1	29.3	17.2	17.1
58889	147.6	60.1	15.2	14.2	45.3	18.6	18.0
68901	312.6	142.2	35.1	27.2.	120.3	42.9	36.1
71377	1345	566.4	145.5	128.6	450.3	175.8	146.5

Table 2 Performance of data service system

Phrase	Web		DwScope		GT-7	
	Internet	LAN	Internet	LAN	Old(LAN)	New(LAN)
Data transfer(s)	18.01	14.01	—	—	12.68	6.15
Decompress(s)	1.92	1.92	—	—	6.52	1.92
Visualization(s)	13.11	5.11	—	—	3.01	0.5
Total time(s)	33.04	21.04	20.54	17.27	22.21	8.55

Table 3 Configuration of computers used for the test

Configuration	Server	LAN client	Internet client
CPU	Xeon 2.8G*2	Intel P3 1.0G	Intel P3 1.0G
Memory	1.0G	256M	256M
OS	Redhat linux9.0	Windows XP	Windows XP

All of above results are tested under single thread. The results show that new system provides satisfied performance. The LZO compression rate amounts to 1:3 for acquiring binary data and the compression speed can be up to 8—18MB/s. Comparing with LZO, MDSplus provides better compression rate at the cost of compression time. Especially for the upgraded GT-7 designed with VC++6.0, its efficiency is almost three times as before. As DwScope runs on the remote desktop mode, its performance is mainly dependent on network bandwidth. At present typical experiments the 3—5 min repetitions of several seconds' plasma discharge are usually encountered. From above parameters, with the help of GT-7 or DwScope, operators have enough time to explore data and control the experiment.

5 Conclusions

HT-7 is different from many other fusion devices such as DIII-D and C-Mod. It is a superconducting Tokamak and it can operate the plasma discharge for a long time as up to 237.86 seconds at present,[8] and so the data service system must correspond to its own characteristics.

In order to meet the needs of researchers to make every kind of physical diagnosis and build virtual fusion laboratory, the authors have built a complete data service system. Firstly, in order to lighten the burden of data acquisition server (the main file server) and lessen the effect from clients, the authors have to look for a reliable way to resolve such problems by separating raw data from analysis data, thus the acquisition server will not be affected by client deeply. It only needs to send the raw data to client when required and needn't do any calculation. Secondly, the authors adopt MDSplus as centralized computer system to manage analysis data. MDSplus is widely used in many fusion institutes, and it has the ability to interface with many commercial data analysis software such as IDL and Matlab. Even if the scientists don't know the data format, how it is organized and where it is stored, they can use standard routines to get data from the server and store the results and comments to MDSplus server. Also MDSplus provides many routines to process data. Thirdly, we just developed a smart and fast data analysis and visualization software GT-7. It has an integrated language, and it provides basic function for processing and analyzing data. Users can easily finish data analysis and visualization only by using mouse and the showing can be scaled and rotated in interactive manner. Lastly, the current system configuration has become an example to establish data system for EAST which is being constructed.

Acknowledgements

The authors thank Z.S. Ji, H.Z. Wang and F. Wang from ASIPP for their critical efforts in developing and supporting the computing infrastructure on which our data service system is based. We would also like to thank T. Fredian and J. Stillerman from Alcator C-Mod and P. Phillips and W.L. Rowan from The University of Texas at Austin for their selfless, timely advice and support of MDSplus. We would also thank L. Zhu and W.Y. Yu for their former design of data analysis and visualization GT-7. Finally we would like to thank the HT-7 group for their valuable feedback during the development process.

References

1 Luo J R, Wang H Z, Ji Z S, *et al*. IEEE Transaction on Nuclear Science, 2002, **492**: 496

2 Documentation, Information available from http://www.oberhumer.com

3 MDSplus Documentation Library, Information available from http://www.mdsplus.org

4 Schachter J, Peng Q, Schissel D P. Fusion Engineering and Design, 2000, **48**: 91-98

5 Davis W, Roney P, Carroll T, *et al*. The use of MDSplus on NSTX at PPPL, March 2002, PPPL-3678

6 Emoto M, Iwadare Y, Nagayama Y. Fusion Engineering and Design, 2004, **71**: 201-205

7 Walters M, Roach C, Fusion Engineering and Design, 2004, **71**: 207-212

8 Achievement in HT-7 Experiments, Information from http://www.ipp.ac.cn

9 Seo Seong-Heon, Hong J S, Kwon M. Fusion Engineering and Design, 2004, **71**: 141-144