A Continuous Data Acquisition System Based on CompactPCI for EAST Tokamak

Feng Wang, *Member, IEEE*, Guiming Li, Shi Li, Yingfei Zhu, and Yong Wang

***Abstract*—For the purpose of long time data acquisition with high speed, a continuous data acquisition system based on the CompactPCI technologies was designed and developed for long time operation of the EAST tokamak. The system can continuously capture data for more than 1000s at 100 k sampling rate and provide the ability of real-time data view and access during the discharge. This paper describes the detailed information of the system including the system structure, workflow, data acquisition card, data transfer and data storage, as well as a discussion of the data access method. The system has been designed and developed and will be demonstrated in the next campaign of the EAST tokamak. The future work is also discussed.**

***Index Terms*—CompactPCI, continuous data acquisition, real time, simultaneous digitizer.**

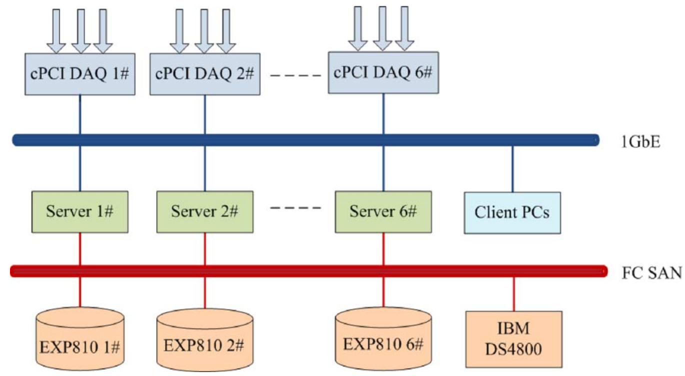

Fig. 1. System structure.

I. Introduction

THE Experimental Advanced Superconducting Tokamak (EAST) is a full-superconducting tokomak device, and one of its objective is to test pulses of up to 1000 seconds with 0.5 MA plasma current. The construction was completed in March 2006 and the first plasma was achieved on September 28, 2006. In February 2007 the device sustained a plasma current of 250 kA for 5 seconds.

The EAST data acquisition (DAQ) system based on Industrial Personal Computer (IPC) was designed and implemented originally for pulse discharge operation in 2006. In the pulse data acquisition mode, data is captured and saved in local memory during the discharge and transferred to the data server after the discharge. In case of this traditional data acquisition system, it cannot acquire data for a long time since the local memory capacity is fixed, for example, on a normal IPC the memory is not more than 4 GBytes. The sampling rate has to be decreased in order to get a longer acquisition time. Furthermore, the system has another disadvantage, which is that acquired data cannot be accessed during discharge, and operators have to wait the data acquisition process to finish. With the increment of the discharge duration, especially in the long-term operation mode, the disadvantages become increasingly intolerable. So a continuous data acquisition system was found required and very necessary for the EAST device. In the continuous DAQ mode data is acquired and transferred to the data server during the discharge, so operator can view data and waveforms at any time during the discharge, even after dynamic change the operation parameters, which is considered important for steady-state operation.

A PC-based continuous data acquisition system has been developed for the EAST magnetic measurements such as flux and magnetic fields with more than 600 channels; however, its maximum sampling rate is not more than 10 k samples per second (KSPS), and it cannot provide simultaneous data acquisition, furthermore PC is not very stable and reliable for continuous data acquisition. These capabilities are unsuitable for the required diagnostics measurements. So a new, higher sampling rate continuous DAQ system is required [1]–[3].

At present, the VersaModule Eurocard (VME) and Compact Peripheral Component Interconnect (CompactPCI) technologies are widely used in large physics experiments including the fusion devices. When compared with the VME bus, the CompactPCI standard is a newer technology and has some benefits such as lower cost, higher bandwidth, and is widely available from several vendors. Thus, the CompactPCI technology was chosen for our new system as well as it is happening for many current control and DAQ systems. The objective of this paper is to present the design and development of a continuous DAQ system based on the CompactPCI technology, which is able to acquire data simultaneously for more than 1000 seconds at 100 KSPS while allowing data to be accessed during the discharge in real-time [4], [5].

II. System Overview

A. System Structure

The new continuous data acquisition system based on CompactPCI for EAST has been designed and developed based on the structure shown in Fig. 1. It contains a set of six DAQ units, each including a CompactPCI computer, a data server and data

Manuscript received May 23, 2009; revised October 19, 2009. Current version published April 14, 2010. This work was supported in part by the National Natural Science Fund of China under Grants 10675128 and a grant from the Innovation Foundation of Chinese Academy of Sciences under Contract KJCX3.SYW.N4.

The authors are with the Institute of Plasma Physics, Chinese Academy of Sciences, Hefei, Anhui 230031, China (e-mail: wangfeng@ ipp.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2009.2037499

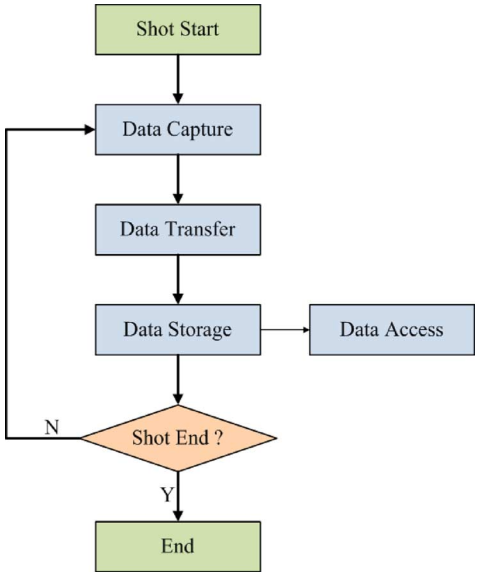

Fig. 2. System workflow.

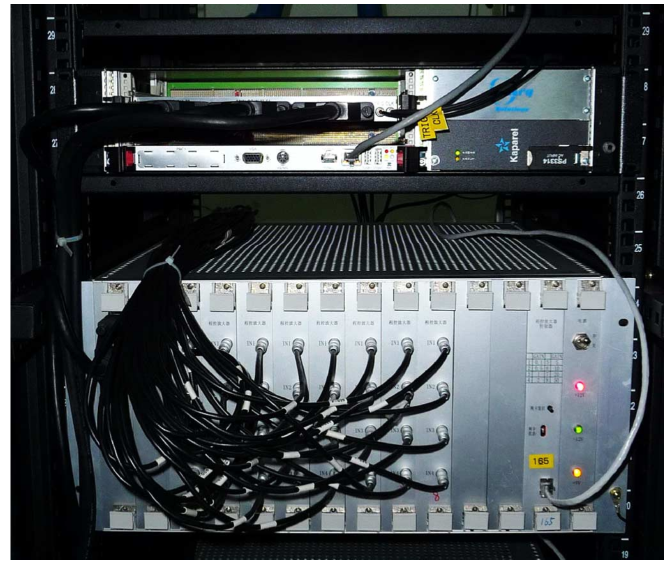

Fig. 3. Hardware of the data acquisition system.

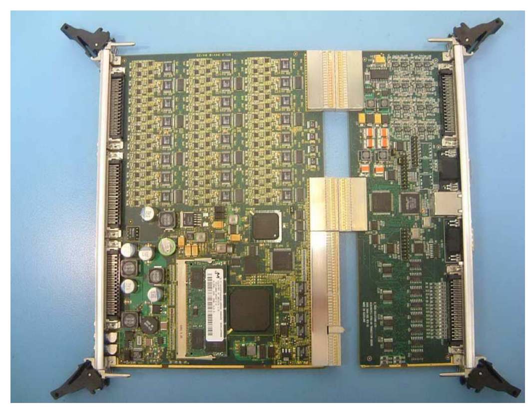

Fig. 4. Simultaneous data acquisition board.

storage units. The system is sub-divided in four parts: data capture, transfer, storage and access. The data capture part includes the six distributed CompactPCI data acquisition computers. The data acquisition network is one of the EAST experimental networks systems, which was chosen to be a full Gigabit Ethernet adopted due to the real-time data transfer capabilities. A large amount of data is produced by the DAQ, so a high performance commercial data storage system was adopted. How to access the data from client computers using the network both in real-time and offline, as fast as possible, is also a very important design requirement for the system. The six data servers provide data transfer and data access services [6], [7].

B. *Workflow*

The main workflow of the system is shown in Fig. 2. Firstly, the diagnostics data are captured and saved into the buffered RAM on board continuously. Secondly, the acquired data is fetched in real-time into the memory of the CompactPCI computer. Thirdly, the fetched data is transferred to the data server plugged to the Ethernet network every 5 seconds, then compressed and saved into the central storage system. Lastly, data can be accessed from client computers. These steps are repeated until the discharge is finished or user aborts the process. The following sections describe the details of these steps.

III. Data Capture

A. *Hardware*

As shown in Fig. 3, one data acquisition unit includes one CompactPCI platform and three signal conditioning crates. The platform is a standard 2U height 6U width four slots crate and complying with the PICMG 2.16 specification. A single board CompactPCI controller (ADLINK cPCI-6840 PM1.6 GHz/2 GB RAM/IDE 160 GB) and a data acquisition card (ACQ196CPCI-96-250) are inserted into the crate. The DAQ card acts as a PCI device of the controller which has three 68 pin SCSI connectors containing each one 32 signals inputs. Each signal conditioning crate has 8 boards with 4 channels. The signals from diagnostics system are firstly input to the conditioning crates and then output to the DAQ card using a standard 68 pin SCSI cable. All crates are connected to the Ethernet network.

B. *Simultaneous Digitizer*

For the implementation of multi-channel, simultaneous data acquisition, an intelligent data acquisition board (ACQ196CPCI-96-250) was chosen for the system; the board was supplied by D-TACQ Solution Ltd. as shown in Fig. 4. The main features of interest to this paper are the following [8]:

1) 16 bit ADC per channel for true simultaneous analog input;
2) high throughput, 250 kSPS/channel;
3) maximum channel density—96 channels in one slot, replaces a rack full of 16 channel data acquisition boards!
4) uses latest silicon for highest performance at lowest cost per channel;
5) deployed in competitive applications down to 1 kHz per channel;
6) true differential input in each channel. Input features high common mode range and input overvoltage protection;

TABLE I
THE ESTIMATED DATA AMOUNT

Sampling rate (KSPS)	10	50	100
Original data / s (MB)	12	60	120
Original total /10s (MB)	120	600	1200
Original total / 1000s (GB)	12	60	120
Compressed data / s (MB)	6	30	60
Compressed total / 10s (MB)	60	300	600
Compressed total /1000s (GB)	6	30	60

7) plant cable interface on the front panel – 3 × SCSI 68 connectors.

C. Software

The Redhat Linux operating system was installed in the CompactPCI controller, where the data acquisition program was developed in C/C++ and compiled using the GNU C tools. In this program multi-threading and socket communications are the key technologies. The multi-thread technology was adopted in order to make full use of the CPU and memory system resources. Socket technology is the API of the TCP/IP networking, widely used for network communication on Linux programs. There are two main threads in the program. Thread-A fetches data from the buffered RAM on board every second. Thread-B transfers the fetched data to data server every 5 seconds.

IV. DATA TRANSFER AND STORAGE

A. Network

The continuous data acquisition system will produce a large amount of data when in steady-state operation; the estimated amount is shown in Table I. When the sampling rate is set to 100 KSPS, the original data size was 120 MBytes per second, while with 1000 seconds of steady-state operation the total size was about 60 GBytes. This is a big challenge for continuous data transfer and data storage. A gigabit Ethernet network was adopted for real-time transfer of a large amount of data as shown in Fig. 1. All the DAQ computers and data servers are connected to the Ethernet.

B. Compression

To decrease the data amount, LZO—a real-time data compression library was adopted. The main features of LZO are as follows:

1) decompression is simple and very fast;
2) requires no memory for decompression;
3) compression is very fast;
4) requires 64 kBytes of memory for compression;
5) allows you to select extra compression at a speed cost in the compressor; the speed of the decompressor is not reduced;
6) compression levels can be set for generating pre-compressed data thereby achieving good competitive compression ratios;
7) includes also a compression level which needs only 8 kB for compression;
8) algorithm is thread safe.
9) Algorithm is lossless.

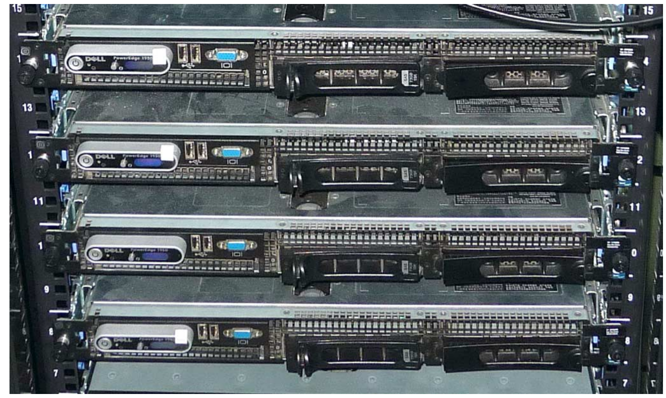

Fig. 5. Data servers for data management.

The LZO compression method has many types of algorithm and parameters; however, we have selected the most common LZO1X-1 since it provides a good balance in compression ratio versus speed, as well as decompression speed. The C version of LZO1X-1 is about 4–5 times faster than the fastest zlib compression level, and it also outperforms other algorithms like LZRW1-A and LZV in both compression ratio and compression speed and decompression speed [9].

Captured data are firstly transferred to the data servers every 5 seconds and then queues up for compression on the data servers. In order to foster the compression speed, a high performance data server (DELL Power Edge 1950) with Quad Core is used for each DAQ computer as shown in Fig. 5. There multi-thread technology can be realized and four files are compressed by the CPU at the same time.

In the test environment, the compression speed was found to be about 50 $\sim$ 100 MBytes/s depending on the different types of signals, while the compression ratio was about 50% of the raw data. On a normal personal computer the decompression speed is about 200 MBytes/s which is approximately three times the compression speed. The data throughput of the system after compression is also shown in the table.

C. Storage

As mentioned above the large amount of data is collected by the data acquisition system continuously, so the data throughput and data volume are both challenges for the data storage system. And stability is also important since during the experiment data system should be on line all the time. Thus, a commercial IBM DS 4800 storage system was chosen due to the proven good performance and wide use in large data systems as well as experiments.

The IBM System Storage DS4800 disk storage system supports a high-performance 4 Gbps Fibre Channel interface for increased host connectivity to deliver necessary bandwidth for high-throughput applications. It is designed for data-intensive applications that demand increased connectivity provided by eight 4 Gbps host channels designed to provide up to 1724 MBps of sustained bandwidth for high-throughput applications through the eight channels directly attached to the host servers or connected to a Fibre Channel Storage Area Network (FCSAN) [10]. There are six EXP 810 expansion

Fig. 6. The data storage expansion units EXP 810.

units with drivers in the system as shown in Fig. 6 and the total storage capacity is more than 50 TBytes.

V. Conclusion and Discussion

With the purpose of continuous data acquisition for long time and steady-state operation of the EAST tokamaks, a new data acquisition system based on the CompactPCI technology has been designed and developed. The system can continuously capture data at 100 KSPS of sampling rate during more than 1000 seconds. A high performance intelligent simultaneous digitizer was adopted which provides 96 channels at 100 KSPS of continuous data capture, and also a commercial storage system used for high throughput data streaming and data management. The system has been tested and will be demonstrated in the next EAST discharge experiment campaign.

Since this DAQ system generates a large amount of data as discussed above, the question of how to access this data easily and quickly is an issue to address. An approach based on timestamp will be used for data management and data access. The main goal of this technology is described as follows [11]:

1) the acquired data is saved into each LZO file every 5 seconds with timestamp for each signal;
2) data is managed and indexed according to timestamp;
3) data can be accessed and browsed based on timestamp;
4) a new 1 K sampled data is extracted from the original data to allow waveform browsing in real-time.

Using this method, the real-time waveform can be browsed during the discharge and all the original data could be accessed after the discharge. A graphical user client tool based on a browser/server approach is being developed for online data browsing.

In future work, we will focus on how to optimize data management and query from large amount of raw data, and on some other data management methods to be investigated such as the MSDplus database [12].

Acknowledgment

The authors would like to thank Dr. P. Milne of D-TACQ Solution, Ltd., for his cooperation and technical support on the data acquisition card. The authors would also like to thank Dr. Y. Liu for her study on data access technology for the EAST data acquisition system.

References

[1] H. Nakanishi, M. Emoto, M. Kojima, M. Ohsuna, and S. Komada, “Object-oriented data handling and OODB operation of LHD mass data acquisition system,” *Fusion Eng. Des.*, vol. 48, pp. 135–142, Aug. 2000.
[2] H. Nakanishi *et al.*, “Steady-state data acquisition method for LHD diagnostics,” *Fusion Eng. Des.*, vol. 66, pp. 827–832, Sep. 2003.
[3] B. Guillerminet and J. How, “The acquisition system for Tore Supra 1000s discharges,” *Fusion Eng. Des.*, vol. 48, pp. 155–161, Aug. 2000.
[4] Y. Buravand and B. Guillerminet, “Status of the Tore Supra continuous data acquisition and retrieval system,” *Fusion Eng. Des.*, vol. 71, pp. 111–115, Jun. 2004.
[5] B. Guillerminet *et al.*, “Tore Supra continuous acquisition system,” *Fusion Eng. Des.*, vol. 60, pp. 427–434, Jun. 2002.
[6] S. Sudo *et al.*, “Control, data acquisition and remote participation for steady-state operation in LHD,” *Fusion Eng. Des.*, vol. 81, pp. 1713–1721, Jul. 2006.
[7] P. Heimann *et al.*, “Status report on the development of the data acquisition system of Wendelstein7-X,” *Fusion Eng. Des.*, vol. 71, pp. 219–224, Jun. 2004.
[8] *“D-TACQ 2G User Guide,”* D-TACQ Solutions Ltd., 2004.
[9] *“LZO—A Real-Time Data Compression Library,”* ver. 2.03, Markus Franz Xaver Johannes Oberhumer, Apr. 30, 2008 [Online]. Available: http://www.oberhumer.com/opensource/lzo/lzodoc.php
[10] *“DS4800 Data Sheet,”* IBM Corporation, 2007.
[11] Y. Liu, J. Luo, G. Li, Y. Zhu, and S. Li, “The EAST distributed data system,” *Fusion Eng. Des.*, vol. 82, pp. 339–343, Jun. 2007.
[12] *“User’s Guide to MDSPlus,”* [Online]. Available: http://www.mdsplus.org