ELSEVIER

Fusion Engineering and Design 43 (1999) 309–319

Fusion Engineering and Design

The TJ-II data acquisition system: an overview

J. Vega *, C. Crémy, E. Sánchez, A. Portas

Asociación EURATOM/CIEMAT para Fusión, CIEMAT. Edificio 66, Avda. Complutense, 22, 28040 Madrid, Spain

Abstract

The data acquisition system for the TJ-II fusion machine has been developed to coordinate actions among the several experimental systems devoted to data capture and storage: instrumentation mainframes (VXI, VME, CAMAC), control systems of diagnostics and a host-centralized database. Connectivity between these elements is achieved through local area networks, which ensure both good connections and system growth capability. Three hundred VXI based digitizer channels have been developed for TJ-II diagnostics. They are completely software programmable and provide signal analog conditioning. In addition, some of them supply a programmable DSP for real time signal processing. Data will be stored in a central server using a special compression technique that allows compaction rates of over 80%. A specific application software has been developed to provide user interface for digitizer programming, signal visualization and data processing during TJ-II discharges. The software is an event based application that can be remotely launched from any X terminal. An authentication mechanism restricts access to authorised users only. © 1999 Elsevier Science S.A. All rights reserved.

Keywords: Instrumentation mainframes; TJ-II diagnostics; Host-centralized database; Signal analog conditioning

1. An introduction to data acquisition in fusion devices

The data acquisition system of a fusion device is an information system which provides resources for collecting, storing, accessing and analysing data. Although data collection is accomplished through suitable acquisition hardware whose configuration parameters can be programmed by a computer system, usually, the acquisition process is controlled by hardware triggers. Actually, a data acquisition system of a fusion machine is made up of several small computer systems (henceforth, subsystems) that can handle different hardware. The variety of subsystems is a direct consequence of the many kinds of measurements that need to be performed. On one hand, it can be difficult, or inefficient, to manage all channels from a single system. On the other hand, special requirements may be necessary for certain diagnostics and, in such cases, dedicated systems must be used. Typically, each subsystem is prepared only for providing temporary storage of data. As a result, it is imperative to have available a data server and to store in it all the information captured from any means. Consequently, the acquisition system of a fusion device must supply some mechanism for integrating, in such data server, the information that is scattered in several subsystems after a discharge.

* Corresponding author. Fax: +34 91 3466124; e-mail: vega@ciemat.es

PII S0920-3796(98)00399-8

Access to the databases is provided by developing software routines that allow the data to be read in a transparent way, hiding from the user aspects such as directories or the exact form in which the data were written. Finally, computational and graphical tools are also provided in order to analyze the discharges. Mathematical, statistical and graphical libraries have been common software packages in the fusion environment, but, nowadays, fourth generation languages are also widely used.

In magnetic confinement pulsed devices, the data acquisition systems present common characteristics. Fig. 1 shows the typical data acquisition cycle which arises as a direct consequence of the pulsed operation mode of the devices, and it is independent of other factors such as hardware instrumentation or device size. The set of systematic activities can be easily identified either in particular data acquisition systems of specific machines, JET [1], TFTR [2], RTP tokamak [3], TJ-I tokamak/TJIU torsatron [4], or in a general data acquisition system like MDS/MDS+ [5]. The first phase, named 'setup', symbolises the set of actions to perform between shots, in order to fit the measurement instrumentation for a suitable signal conditioning and to program the acquisition channels. The next step, 'discharge', actually is made of two parts. Firstly, it begins the 'preshot' activity, whose aim is to arm all acquisition hardware and in which all software tasks related to data capture are suspended, giving the control of the acquisition to hardware triggers. Secondly, when the discharge takes place, every measurement channel receives information and provides local storage for it.

After finishing the discharge, the stage called 'preanalysis' begins. Having access to the data residing in several subsystems, it is possible to perform a first analysis of the signals stored there, such as, in the most simple case, the visual inspection of the signals captured. This preanalysis may be sufficient to determine the quality not only of the discharge, but of the signal conditioning as well. From these observations, it is feasible to give feed back the first phase of the cycle, and to configure better parameters for the next shot for both, device control and data acquisition.

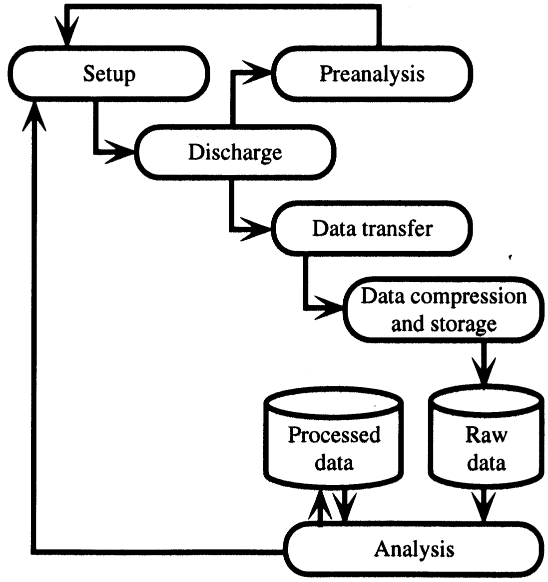

Fig. 1. Typical data acquisition cycle.

In parallel with the actions just mentioned above, the 'data transfer' step also begins. Due to the fact that the different subsystems usually are not interrelated themselves, in order to be able to analyze data captured on a given subsystem, it is necessary to have access to it. But, except for a

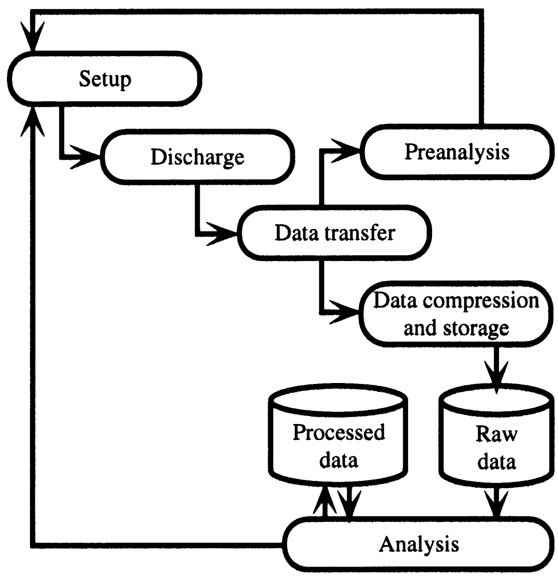

Fig. 2. TJ-II data acquisition cycle.

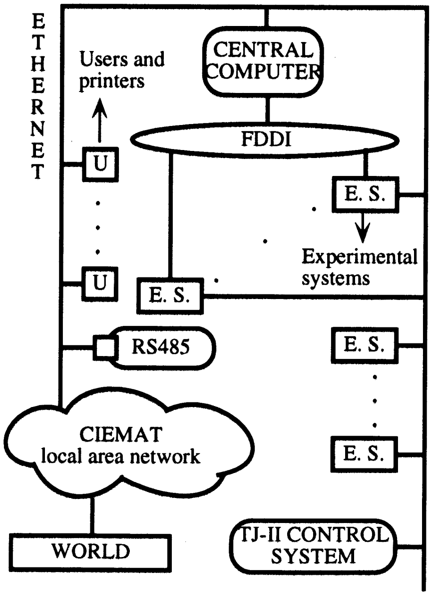

Fig. 3. Structure of networks for TJ-II.

simple data visualisation, it is not possible to perform a simultaneous analysis with signals acquired on different subsystems. To do that, data must be integrated under a single device and for this reason, all signals are transferred to a data server. However, the transmission must be finished between shots and in the least possible time. This is a consequence of the very limited storage space of the acquisition hardware (generally, digitizers with only tens of Kbytes). The data of a shot are rewritten over the data of the previous discharge and, therefore, in order to avoid the loss of information, the data must reach the data server between consecutive shots.

To perform computations with any one of the data collected, it is required to generate the 'discharge file' in the data server as fast as possible. This file is created in the phase 'data compression and storage' and keeps all the information related to a discharge. The 'discharge file' allows unification of the data writing in a common format and with the same structure, independently of instrumentation of capture or acquisition parameters among other factors. In this stage, in order to save storage space, the information is also compacted.

When the 'discharge file' has been built, the last step of the cycle, 'analysis' begins. Each diagnostician analyzes his data according to the proper codes to attain plasma parameters. Generally, these analyses read raw data, process them and feed databases which contain data with physical sense. The raw data are not modified, because the processed data are susceptible to changes. Of course, there is no temporal limit to perform calculations. However, sometimes it is important to do some computations in order to give feed back to the first phase of the cycle between discharges.

Nowadays, acquisition of tens of Gbytes per month is very common for large and medium-size fusion devices which only allows a few months of data to be stored on disk. To free disk space in order to store new discharges, data are migrated periodically to a cheaper storage media, typically magnetic tapes. Big tape libraries are produced

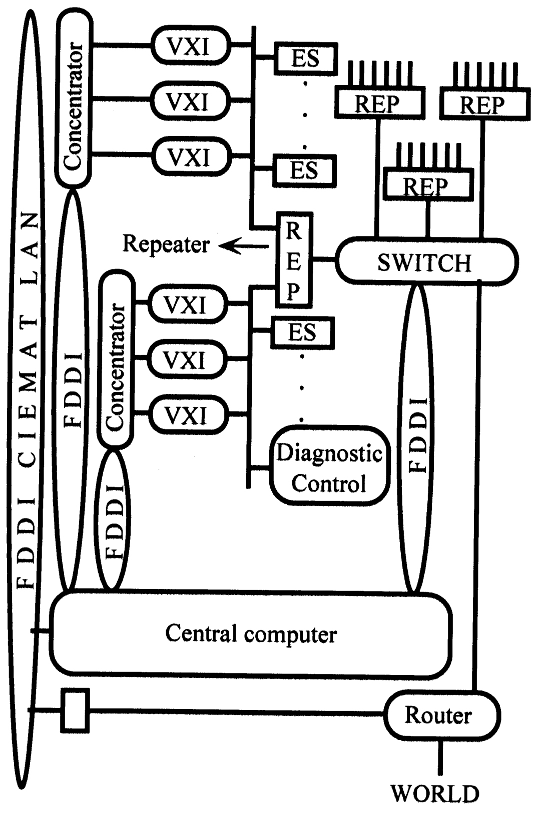

Fig. 4. Networks developed for TJ-II.

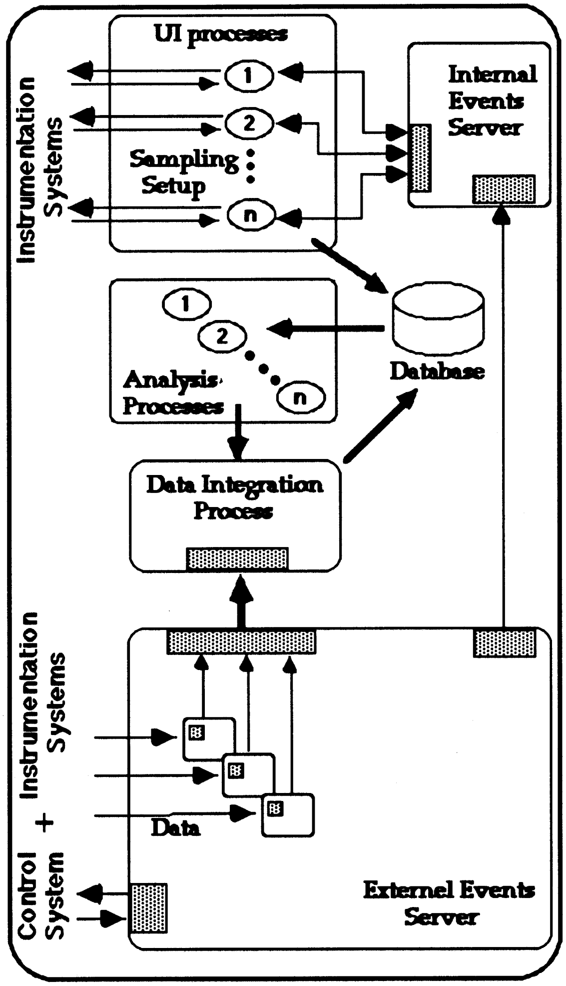

Fig. 5. Software architecture of the TJ-II DAS application.

and they usually work in an automatic manner since it is important to be able to access a large portion of old data automatically, because data retrieving can be demanded frequently.

According to the precedent paragraphs, four critical factors appear in the data acquisition system of a fusion device:

- time for accessing the first data after a discharge.
- time to transfer all data of a shot to the data server.
- time to create the 'discharge file'.
- storage capability.

The label 'critical' arises as a consequence of the influence that those factors can have in the normal operation of the devices. It is absolutely essential that the data acquisition system is adapted to the way of operating a fusion machine, and the inverse condition must never happen. For instance, a device has its shot repetition period fixed by the construction parameters of the machine (coils, power supply, cooling system, etc.). This means that the data acquisition has to provide response times less than such period in every point related to discharge production.

In general, all data acquisition systems of fusion devices follow the cycle shown in Fig. 1. The difference among them lies in the particular solutions chosen to solve the critical aspects mentioned above.

2. The data acquisition system for the TJ-II heliac

The TJ-II heliac is a medium size helical axis stellarator that is located at CIEMAT [6]. It is a four period stellarator with major radius 1.5 m, average plasma radius between 0.1 and 0.2 m, and nominal toroidal field 1 T. It has been designed to study the effect of the helicoidal magnetic axis on plasma confinement for different values of rotational transform. In any configuration, the plasma will have a beanshaped magnetic structure that rotates poloidally when one moves in the toroidal direction. Discharges will be 0.5 s long.

The starting point for beginning the design of the TJ-II data acquisition system was to define five general conditions. The system had to be open, meaning that only standards of instrumentation, computing or communication must be considered. This characteristic is stated in order to ensure continuity and maintenance. Secondly, the system must be flexible, i.e. able to integrate and manage different technologies with respect to physical equipment. In this way, the system is not limited to use a unique platform to solve the requirements of acquisition. Another property was modularity. The system had to be able to provide a wide range of solutions within a common framework. It implies a recommendation for

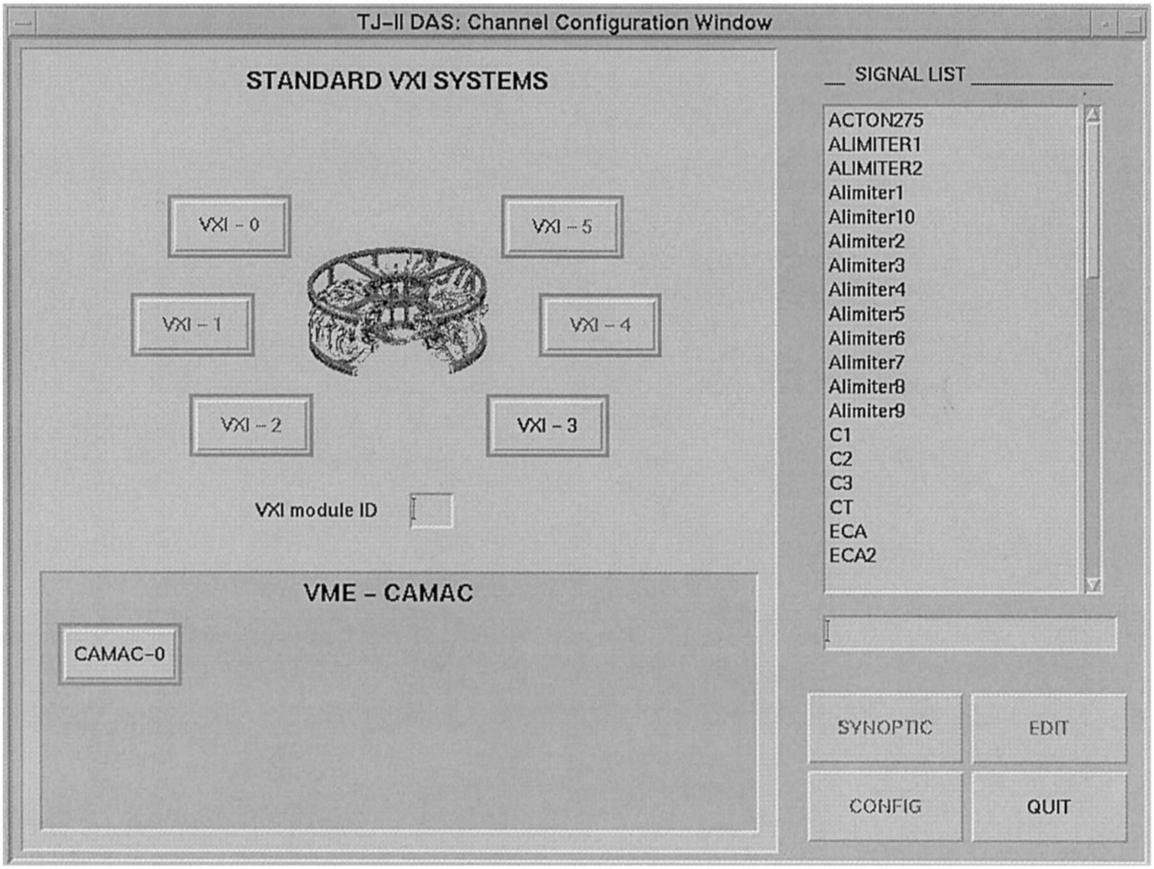

Fig. 6. View of the synoptic mode.

the use of modular instrumentation, which allows coverage of a broad band of requirements while maintaining the reference environment. Also, it was necessary to support growth capability, which is a very important requirement in any data acquisition system. Finally, the system had to show certain fault tolerance, in order to minimize the impact of faults in the daily work routine. To this end, it is necessary to put redundant elements where possible.

After these basic principles, it was necessary to establish others that took into account the general characteristics of the fusion environment and the particular ones of the TJ-II. The new system had to provide the following capabilities in the most automatic, safe, fast and efficient possible way:

- to program measurement channels.
- to acquire data.
- enough short-term and long-term storage.
- management of databases.
- software libraries for data accessing.

One important capability, that usually is not integrated in an acquisition system, is signal conditioning. This compels the insertion of proper instrumentation before the channel input, which implies more physical equipment. In order to minimize the instrumentation, it was decided to incorporate the signal conditioning in the general measurement channels to be developed for TJ-II.

In order to be able to program the experimental systems (equipment for data acquisition or diagnostic control systems) remotely, the condition that they be programmable through software was imposed. It was also established that the channels were distributed around the TJ-II, in different locations. In this way, the channels can be located

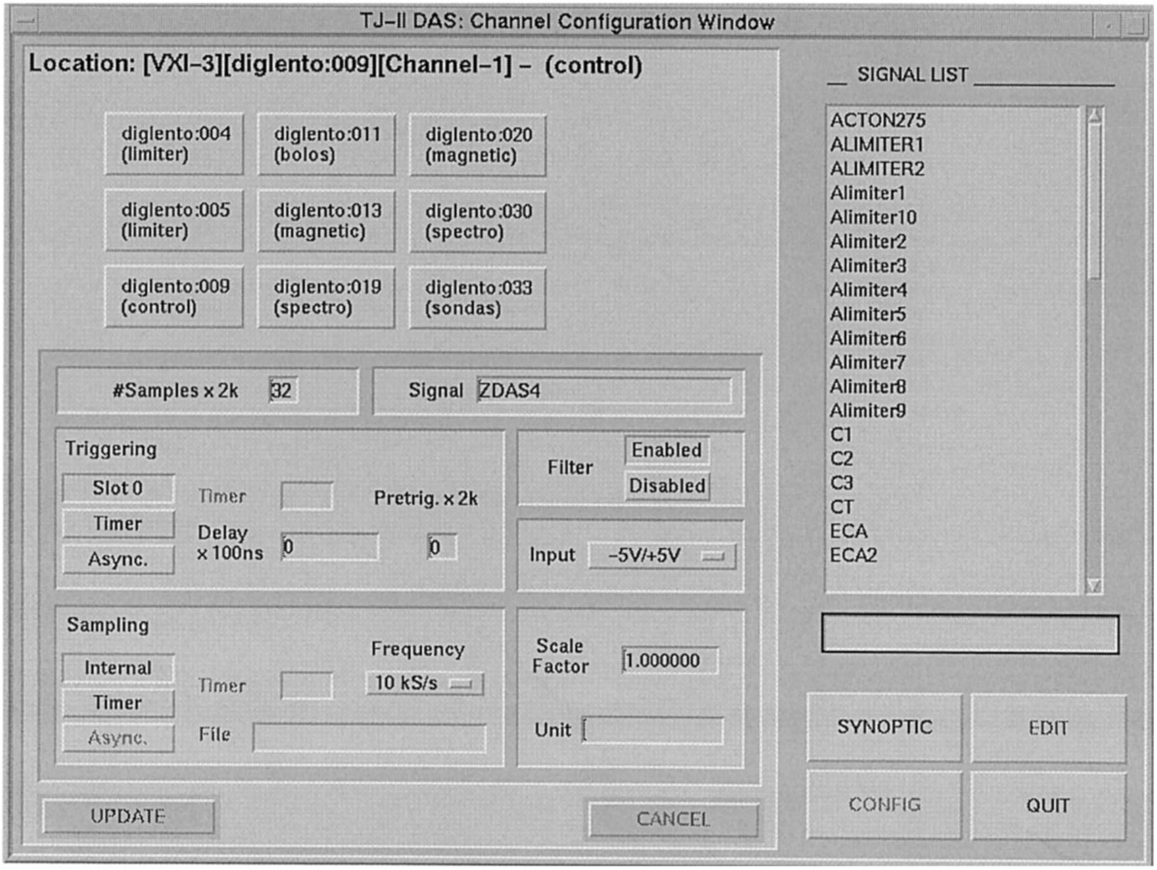

Fig. 7. View of the configuration mode.

near the diagnostics and analog signals will travel through short lengths of cable, which is fundamental in the very hostile fusion environment from the electromagnetic point of view.

Every experimental system in TJ-II will have processing capabilities. Connectivity among them will be achieved by means of local area networks and, in addition, the experimental networks will have to be linked to the CIEMAT local area network.

The frequency of discharge repetition determines a characteristic time in the operation cycle of each fusion device. By design, the TJ-II shots will be repeated every 5 min. This period bounds the time for transferring data from the subsystems to the data server. Typical transfer time in fusion devices is on the order of minutes, which was considered very long for the TJ-II requirements. In order to reach times of tens of seconds, the cycle shown in Fig. 1 has been modified for the TJ-II data acquisition. In the TJ-II cycle, instead of beginning the 'preanalysis' phase just after 'discharge' and in parallel with 'data transfer', 'preanalysis' will be executed after ending 'data transfer' and simultaneously with 'data compression and storage' (Fig. 2). The data will arrive to the data server and will be put, without compressing, in temporary storage. From that moment, all data will be available for any process. The processes related to 'data compression and storage' will access to the information as any other user's process. Once the data have been compacted, the temporary storage will be deleted.

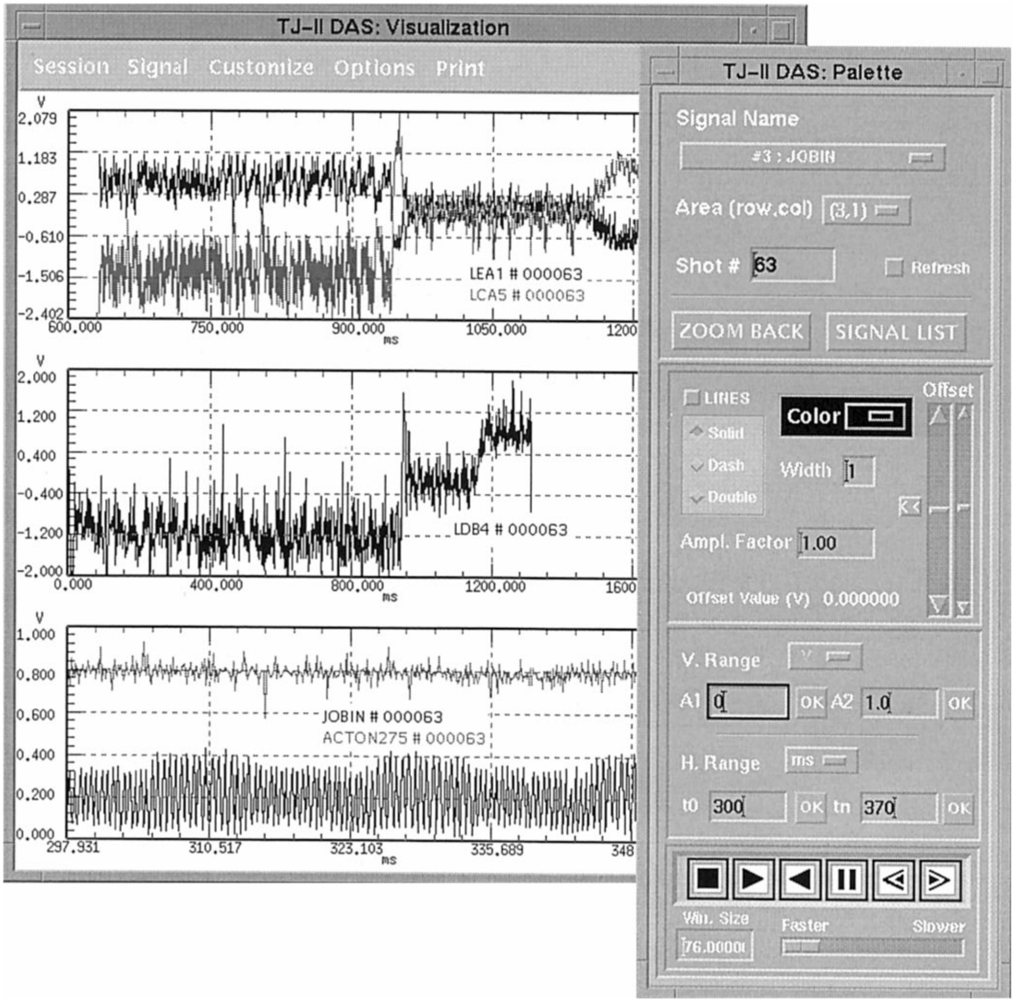

Fig. 8. View of the visualisation mode.

3. System elements

The architecture of the TJ-II data acquisition system was designed to satisfy all requirements mentioned above. The system can be described as a set of different elements which, acting in a coordinated way, reach the desired aims. These elements can be grouped in two general types. On one hand, there exist a set of hardware resources that provide capabilities for data taking (measurement channels), storage (central server) and, connectivity (local area networks). On the other hand, we have developed software resources for, firstly, minimizing, as much as possible the impact of big requirements of storage. Secondly, providing tools for data management and retrieving, and finally, programming the hardware in an efficient and orderly way.

3.1. Hardware resources: measurement channels

The TJ-II data acquisition system is going to provide general measurement channels that can coexist with other systems of data collection (CAMAC, VME, dedicated systems, etc.). Initially, these channels will be transient recorders. The fastest ones perform at conversion frequencies up to 20 M samples s^{-1}. They are based on the VXI standard and have been designed and developed at CIEMAT. VXI was chosen because it ensures very good characteristics in relation to electromagnetic and radio-frequency shielding, power supply, cooling, components integration, development possibilities and on-line processing capabilities.

Three hundred channels were estimated as sufficient to achieve a versatile and general purpose data acquisition system, adapted to the initial requirements of the TJ-II device. A full description of them can be found in [7]. Several features were added to the basic analog/digital conversion, in order to optimize the treatment of input signals: single ended or differential inputs, programmable analog signal conditioning (amplification and low pass antialiasing filter), 12 bits of resolution, up to 1 MByte of memory, trigger delay, pretrigger samples storage capability and offset correction. Finally, some channels have a digital signal processor (DSP) that can be programmed for real time processing of the samples. This component provides a great flexibility in computations and allows a faster analysis of the discharges. In addition to this, two important benefits are achieved for the TJ-II modules making use of DSP's together with the oversampling technique: low order filters are used for the antialiasing low pass filters (in spite of the cut-off frequency and sharpness of the transition band) and the signal to noise ratio is increased some additional bits of resolution.

A real time operating system runs in the VXI mainframes, ensuring a bounded response time to events.

3.2. Hardware resources: central computer

The central computer of the TJ-II data acquisition system is the key element of the system. This computer supervises and coordinates the actions between subsystems. It receives experimental data and feeds the device database [8]. The server is a multitasking computer, with a high input/output bandwidth, running UNIX operating system. The storage capacity for TJ-II data is 125 Gbytes. Data will be protected by using the RAID 1 technology up to the moment in which the information will reside on magnetic tape. The low efficiency of RAID 1 in writing is compensated by using intelligent controllers with cache memories of 64 MB. These controllers satisfy a writing operation when the data are stored in the cache, without waiting for the real writing on disk. This fact speeds up significantly the execution of codes.

3.3. Hardware resources: local area networks

The TJ-II data acquisition system can be seen as a set of scattered subsystems (with processing capabilities and remote programmability) all of them coordinated by a supervising system that will keep and manage the device databases. The common links among experimental systems will be the local area networks, which provide high transmission speed in a safe manner. Due to the different types of control devices for subsystems (PC's, workstations, computers, modular instrumentation, etc.) running several operating systems (DOS, real time, sharing time, etc.) the LAN's ensure the connectivity among such systems by using a unified network protocol like TCP/TP.

Fig. 3 shows a block diagram of the architecture, although the actual development has not been so simple, as can be seen in Fig. 4.

Three different LAN's will be used for TJ-II [9]. Firstly, an ETHERNET network covers the basic requirements of connectivity between subsystems. Such a LAN provides interactive connections of users (U) with the central computer, either for programming measurements during the TJ-II operation or for accessing data. The access to the databases must be carried out not only from nodes of this network, but from anywhere in CIEMAT and in the world.

The second LAN is a high speed network (FDDI) that is used for fast data transferring after a discharge. The requirement of high speed is a consequence of having to finish in the least possible time the 'data transfer' phase of the TJ-II cycle. Those subsystems that can not be linked to this second LAN, will transfer their data through the ETHERNET network.

Some diagnostics may require an automated control of some parts like vacuum systems, engines, mechanical arms, valves, etc. These systems will be linked through a LAN in order to be handled remotely. Some of them will need real time communications and, therefore, the LAN must satisfy this requirement. But many of the diagnostics will be controlled by means of programmable logic controller devices (PLCs) not supporting the TCP/IP protocol. To overcome this problem, a third LAN will be in operation. This network will be a RS-485 LAN, which is a very suitable network for connecting control equipment in an easy and efficient way. The RS-485 network will be linked with the ETHERNET network.

3.4. Software resources: data compression

If the data acquisition in a TJ-II discharge were performed only with the 300 VXI channels and the channels were programmed to fill completely the memories, more than 35 Gbytes per month of storage would be necessary. It is clear that this vast quantity must be reduced. But such a reduction must not be accomplished at the expense of losing data. Thus some form of data compression is required to reduce storage.

The method to be used for the TJ-II data is an encoding technique which allows the signals to be stored without distortion that is based on delta compression (i.e. to store the differences between the digital codes of adjacent signal samples instead of storing the digital codes). The technique is a three step algorithm [10]. The first step consists of a delta calculation. The resultant deltas are then encoded according to a variable-length bit allocation (satisfying a prefix code property), which is signal independent. Finally the resulting codes are stored in tandem. Tests on the technique show compression rates over 80% with negligible computation time.

3.5. Software resources: data management and retrieving

Data retrieving is accomplished through the use of software libraries specially developed to this end. But the management of the TJ-II database implies other important tasks. These tasks must be carried out in an automated manner by means of computer processes also running in the TJ-II central server. Typical operations related to this point are: automated integration of new data coming from measurement instrumentation or analysis codes, backups, data access, transfer of data from quick access magnetic media to an automated slow access mass media, handling of the transfer criteria based on the signal age and priority, control of data location, and retrieval of data to fast access media when required.

3.6. Software resources: application software

The TJ-II DAS software has been developed to provide the experimental staff with access to TJ-II DAS instrumentation modules, diagnostic control systems and discharge database, through a window-oriented user interface (UI). The UI program is a host-centralized application, running over UNIX with the X/OSF motif environment that is aimed at being executed in parallel by a determined number of registered users (typically 30).

In operation time, several data resources need to be available to the UI. Signal records of the last shot are temporally stored in shared memory. The rest of the signal records are compressed and saved into hard disks and magnetic bands. An exhaustive mapping of the instrumentation systems and modules is maintained in shared memory, and updated when a change occurs. The list of all the signals that are handled by the application, as well as their associated information, are also available from shared memory. This list is updated when some signal information has been modified or a new signal has been created.

One main challenge for the UI is to update its display in real time. In order to reach this objective, several synchronization mechanisms have been implemented as represented in Fig. 5. The external events server (EES) is a UNIX daemon process running during TJ-II operation. This process is designed to handle every query coming from the diagnostic and control systems ('external world'). Examples of queries may be the notification of mainframe boot or data arrival. When a VXI mainframe is switched-on, a boot program gets information about the detected devices and sends it to the EES, which will update the system mapping.

The internal events server (IES) is a TJ-II operation daemon. This process is dedicated to handle events that come either from the EES process or from the internal world, i.e. the processes and resources localized in the main host, It is responsible for dispatching events, like user login, system boot and signal creation to all the UI processes, thus allowing them to refresh their display. The IES process is based on the X display management control protocol (XDMCP), where users enter their name and password, either from a X terminal or a UNIX workstation supporting this protocol. This process handles users' logins and logouts ensuring a single session for each user. For this purpose it maintains a record of the logged-in users' names and their displays in its memory data space. Other features of the application that need information about users logged-in, e.g. the mail service, can get this information by establishing a connection with the EES process. Communication with these servers is established through FIFOs and BSD sockets and follows a specific protocol.

As a result of the login procedure, the user interface process is started and displays two windows, respectively at the top and the bottom of screen. The former is the 'task bar', which gives access to the features of the interface. The latter is the 'message bar', where the DAS application displays the current shot number and information messages like a new shot arrival. The UI's thread of execution can be in one of the following operation modes: standby, synoptic, channel configuration, signal edition and visualization.

Synoptic mode displays the set of available instrumentation systems (VXI, VME or CAMAC) by means of visual controls and an exhaustive list of signal identifiers (Fig. 6). By a simple click over a motif pushbutton the UI process enters the configuration mode where module and channel setup of the associated system are displayed and may be changed by simple point and click operations (Fig. 7).

Transition to the signal edition mode is performed by a simple click over the signal list. This mode allows the edition of signal characteristics such as type (time history, profile, etc.), owner user name and group, storage priority level, experimental parameters and comments.

The visualization mode (Fig. 8) provides fast data visualization tools for signal records, including y versus x plot, 2D profiles, 3 and 4D plots (through integration with the PV_WAVE software package), zoom operations, and an animation facility. Automatic plot refresh at new shot arrival is also possible on a persignal basis. The visualization mode also includes standard signal processing facilities, (i.e. spectral analysis, transforms, filters and statistical processing) and allows the link with external code for specialized analysis.

In addition to the operation modes described, the application provides a dedicated mail service for fast information interchange between users during TJ-II operation.

4. Summary

The TJ-II data acquisition system is a general system that can integrate multiple experimental systems (data capture and diagnostic control). All of them must have processing capabilities and will be linked, among themselves and with the central server, through local area networks. In operation of the TJ-II, the systems can work independently of the central server or execute software in agreement with the software running in the central computer. In the first option, users must open a connection with the systems and, after a discharge, the data integration in the databases is not automated. In the second one, users open only one session in the central computer and can operate the experimental systems from a single computer session. To this end, a special shell program, based on X window/motif environment, will provide user interface. Data transfer to databases is automated. Communications between experimental systems and the central server will be carried out by means of the TCP/IP protocol.

Other resources provided by the central server are graphical and computational tools for data analysis, data compression, off-line access to databases, reception of data and its integration in the databases at any moment and management of the databases.

General transient recorders based on the VXI standard have been developed at CIEMAT for the TJ-II. Among their characteristics, we can emphasize the analog signal conditioning (amplification and antialiasing low-pass filter) and the real time signal processing making use of a DSP. The memory per channel can be 1 Mbyte and the sampling rate can reach 20 M samples s^{-1}. The channels have a resolution of 12 bits and they incorporate a very flexible trigger protocol that allows acquisition of a variable number of pre-trigger samples. In addition, all configurable parameters of the channels are software programmable.

References

[1] H. Van der Beken, H.E. Clarke, R.F. Herzog, et al., IEEE Trans. Nucl. Sci. 36(5) October 1989, p. 1639.
[2] N.R. Sauthoff, R.E. Daniels, The PPL Computer Division, Rev. Sci. Instr. 56(5) May 1985, p. 963.
[3] P.C. Van Haren, F. Wijnoltz, Comput. Exp. Phys. 7(6) Nov/Dec 1993, p. 638.
[4] L. Pacios, J. Guasp, J. Vega, Rev. Sci. Instr. 63(10) Oct 1992, p. 4806.
[5] G. Flor, G. Manduchi, T.W. Fredian, et al., Proceedings of the 7th Conference Real Time 91, 1991, p. 109.
[6] C. Alejaldre, J. Alonso, J. Botija, et al., Fusion Technol. 17 (1990) 131.
[7] J.L. De Pablos, P. Olmos, J. Vega, et al., IEEE Trans. Nucl. Sci. 43(1) Feb 1996. p. 229.
[8] J. Vega, C. Crémy, E. Sánchez, et al., Rev. Sci. Instr. 68(1) Jan 1997, p.959.
[9] J. Vega, A. Mollinedo, A. López, et al., Rev. Sci. Instr. 68(1) Jan 1997, p. 963.
[10] J. Vega, C. Crémy, E. Sánchez, et al., Rev. Sci. Instr. 67(12) Dec 1996, p. 4154.