Fusion Engineering and Design 84 (2009) 2136–2144

Contents lists available at ScienceDirect

Fusion Engineering and Design

journal homepage: www.elsevier.com/locate/fusengdes

New developments at JET in diagnostics, real-time control, data acquisition and information retrieval with potential application to ITER

J. Vega[a,*], A. Murari[b], B. Carvalho[c], G. de Arcas[d], R. Felton[e], M. Riva[f], M. Ruiz[d], J. Svensson[g], JET-EFDA Contributors[1]

[a] Asociación EURATOM/CIEMAT para Fusión, Avda. Complutense, 22, 28040 Madrid, Spain
[b] Associazione EURATOM-ENEA per la Fusione, Consorzio RFX, 4-35127 Padova, Italy
[c] CFN, Associacao IST/EURATOM, 1049-001 Lisboa, Portugal
[d] Grupo de Investigación en Instrumentación y Acústica Aplicada, Universidad Politécnica de Madrid, Crta. Valencia Km 7, 28031, Madrid, Spain
[e] EURATOM/UKAEA Fusion Association, Culham Science Centre, Abingdon, UK
[f] Associazione EURATOM-ENEA sulla Fusione, C.R. Frascati, C.P. 65, I-00044 Frascati, Italy
[g] Max-Planck-Institut für Plasmaphysik, Teilinstitut Greifswald, EURATOM Association, Wendelsteinstr. 1, 17491 Greifswald, Germany

ARTICLE INFO

Article history:
Received 4 November 2008
Received in revised form 9 February 2009
Accepted 10 February 2009
Available online 7 May 2009

Keywords:
Real-time control
Data acquisition
Information retrieval
Plasma diagnostics

ABSTRACT

In magnetic confinement fusion, the operation of next generation devices will be significantly different compared to present day machines. The duration length of the discharges will require abandoning the traditional paradigm of processing and storing the data after the shot. In fact most information will have to be made available in real-time. The significant issues of machine protection will require more sophisticated and at the same time more robust feedback control schemes. Another very important issue emerged in the last years of JET operation, and which is expected to become more severe in ITER, is the large amount of data to be analysed, which cannot be handled in the most efficient way with traditional methods.

In order to prepare for the operation of ITER, some tests are being performed at JET. The capacity of the real-time network has increased in the last years, and many more systems, mainly diagnostics have been connected to it in order to test their reliability and to assess the quality of the information they can provide for feedback control. To reduce the amount of data, a prototype of real-time adaptive data acquisition techniques is being implemented, to adjust the acquisition frequency to the time resolution of the phenomena to be analysed in the plasma. Lossless data compression techniques have been refined and various intelligent signal processing methods have already been implemented to allow an easier and more objective first screening of the data. To allow scientists from wide and diffuse communities to participate in the scientific and technical programme, various innovative tools for remote participation and experimentation are also being actively investigated.

1. Handling of measurement data in reactor grade fusion devices: the problem in perspective

Magnetic confinement fusion was originally investigated in small devices, with relatively low energy content, operated in pulsed mode. These small machines were very prolific in terms of scientific results but revealed clear scaling laws, according to which the plasma volume is a key ingredient in achieving parameters of reactor relevance. The positive trend of plasma performance with volume forced, in a certain sense, the scientific community to increase the dimensions of the devices, in order to improve efficiency and in perspective converge on machines capable of producing electricity at competitive costs. In order to achieve higher efficiency and reduce the stress of the materials, continuous operation is also considered very desirable, if not essential, for a viable fusion reactor.

The need to increase the energetic content of the plasma and the length of the pulses are essential ingredients in this quest for an alternative energy source but they certainly pose significant scientific and technological issues. Some of the problems have been studied for many years. In particular the severity of plasma wall interactions has already motivated both plasma physics and material technology activities. In most major devices, significant efforts are being devoted to develop plasma scenarios with more benign edge conditions, whereas new materials are being tested for the plasma facing components. In this framework, the adoption of tung-

* Corresponding author. Tel.: +34913466474.
E-mail address: jesus.vega@ciemat.es (J. Vega).

[1] See the appendix of F. Romanelli et al. Proc. 22nd IAEA Fusion Energy Conference, Geneva, Switzerland, 2008.

0920-3796/$ – see front matter
doi:10.1016/j.fusengdes.2009.02.055

sten (W) in Asdex Upgrade and the future installation of a Be wall with a W divertor in JET are big projects [1,2], which reveal the determination of the European community to attack this issue. The higher plasma energy content, the increased length of the pulses and the less forgiving wall materials all pose additional issues also for diagnostics, data acquisition and real-time control.

The aforementioned additional difficulties have been recently experienced on JET. First of all, the pursuit of more performing plasma scenarios has motivated the measurements of more quantities, with higher spatial and time resolution. As an example, JET has already shown the capability to acquire more than 10GB of data per shot with its new set of diagnostics, to be compared with 243KB for the first JET discharge. This trend is estimated to continue at a rhythm compatible with Moore's Law, as already experienced also by smaller devices like TJ-II. Moreover, the need to push the boundary of the operational space normally imposes higher risks, in terms both of anomalous events, like disruptions, and thermal loads on the first wall. This requires significant more stringent demands on machine protection and, consequently, diagnostics also have to provide more data in a reliable way to implement the necessary control strategies. The need to find a good trade-off between improvement of plasma performance and safe operation of the device will require a significant rise in the amount of measurements to be acquired both in JET, for the operation with the new wall, and ITER. For example just the infrared cameras for surveillance are estimated to produce in ITER almost 2TB of data for a 10 min discharge. This amount of information, sometimes to be analysed on a relatively short time scale to provide input to the following discharge, cannot be treated with present day methods.

Another trend of current research, which is increasingly showing the inadequacy of present day solutions, is the emphasis on non inductive operation, to extend the discharge duration. In addition to requiring the acquisition of data for longer periods, this so called advanced Tokamak programme normally relies on quite sophisticated real-time control of several plasma quantities. To cope with these needs, JET real-time network has now grown to include about 30 systems and 500 signals.

For these and other reasons, described in more detail later, the organization of the entire machine operation during the experiments, adopted in small and medium size machines and illustrated in Fig. 1, is not considered adequate for the next generation of devices. In this traditional approach the parameters of the next discharge are set up in feed-forward, on the basis of off-line data analysis performed on previous similar experiments. The predefined measurements are then collected locally by the various diagnostics and sent to a central database after the shot for further analysis. In the perspective of next generation machines, this approach is insufficient in many respects and a new paradigm will be necessary. The main new elements in the general organization of the experiments are illustrated in Fig. 2. First of all provisions will have to be made for a centralised Event Manager, in control of the available actuators and capable of intervening on the systems in real-time. This will require a series of real-time tools, from data compression to fast algorithms, to distribute and analyse the data, so that the plasma status can be interpreted quickly and accurately enough to identify the proper strategy for intervention (see Section 2). Since a much more extensive use of images, provided by various infrared and visible cameras, is expected in ITER and also in JET during next years, coherent policies to reduce the amount of data to be stored will become useful (if not necessary for the distribution of information). In addition to data compression techniques, event oriented measurements methods will have to be developed for ITER. At JET a prototype of adaptive frequency acquisition techniques is being implemented, to increase the data sampling only when interesting phenomena are present (see Section 3). In addition to the infrastructures required by the feedback control of the plasma and efficient data acquisition, the enormous amount of data generated will have to be handled in a different way. The problem of information retrieval is already very evident in present day machines, where very often only a limited amount of the data is properly processed for lack of automated tools to retrieve interesting physical events (see Section 4). Since in Fusion the shape of the signals carries a lot of information about the occurrence and nature of the events to be studied, the approach of structural pattern recognition has proved to be particularly effective in the first screening of vast amounts of data. Again, in the line of supporting a more general use of cameras, these methods are being extended also to images.

The fast developments in computers science, both hardware and software, allow now various attractive forms of remote participation. Indeed ITER is expected to provide adequate facilities for this purpose to guarantee the proper involvement of all the partners. This trend is already very evident in JET, which is now visited by hundreds of scientists from all the National Associations. The most advanced stage of 'remote experimentation', where scientists not on site will be granted a certain level of control of part of the

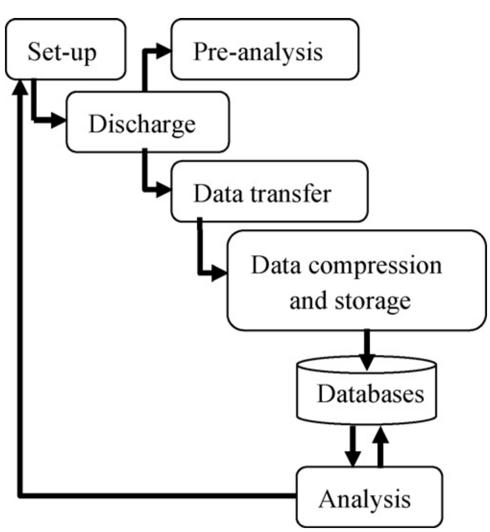

Fig. 1. Traditional organization of the experiment.

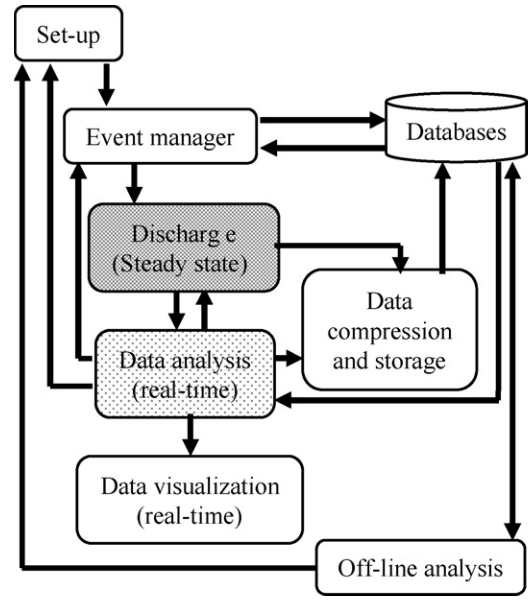

Fig. 2. Conceptual framework for a more advanced organization of the experiment.

machines, has to be looked in the long term perspective (see Section 5).

2. Diagnostics and real-time

In the next generation of devices, several diagnostics will play a different role than in present day machines. Given the complexity and potential risks of hotter plasmas, with higher energy content and longer durations, many measurements will be indispensable to allow safe operation of the devices. In addition to providing the necessary information for the physical studies, many diagnostics will have therefore to guarantee real-time data of enough quality on every shot. These requirements will have implications in the design of the systems, which will have to be in general more reliable than in present day experiments. Many diagnostics will have also to provide processed data in real-time, to allow the implementation of physical based feedback control strategies. In addition to more reliable hardware, more robust algorithms are therefore to be developed in order to process the information on the time scales needed. Moreover in situ periodic calibration techniques will have in some cases to be developed in order to overcome drifts and instabilities of the measurement systems.

At JET, in the last years, the need of a more consistent set of real-time signals has emerged very clearly. To ensure both the safe operation of the device and adequate support to the experimental programme, progressively more systems have been connected to the real-time network. Today the real-time needs are managed by the so called Real-Time Measurement And Control Facility, which consists of Plasma Diagnostics, Plasma Heating and Fuelling systems, a Signal Server (RTSS), a high-level control programming environment Real-Time Central Controller (RTCC) and a data communications network [3]. The Signal Server acquires analogue signals from some Plasma Diagnostics and digital data-sets from other Diagnostics and the heating/fuelling systems. The Central Controller runs user-defined control scripts and sends control messages to the Heating and Fuelling Local Managers, which implement the required actions. Thus, the plasma experiment can adapt to the state of the plasma in real-time. This infrastructure, shown graphically in Fig. 3, supports a wide range of applications, from simple protections based on a threshold to sophisticated model based feedback schemes. These applications, which have already been systematically used in the past campaigns, can be typically grouped as follows:

(1) Plasmas/machine condition protection e.g.:
   - Step-down heating power when the neutron yield exceeds some threshold to avoid pressure-gradient-driven disruption.
   - Inhibit pellet momentarily to avoid density-driven disruption.

(2) Event-driven experiments e.g.:
   - Feed-forward control of heating when magnetohydrodynamic (MHD) mode signals reach thresholds.

(3) Steady state experiments e.g.:
   - Single Input, Single Output (SISO) feed-back control of diamagnetic energy (W_{dia}), poloidal beta (β_{p}), total radiated power, He concentration, internal transport barrier (ITB) criteria.
   - Multiple Input, Multiple Output (MIMO) control of q-profile.

The original network (c.1995) used analogue signals and ethernet. Most of the analogue signals were often "raw" instrumental signals, rather than physics signals, and were not very useful. The ethernet traffic had to be carefully timed to avoid collisions and loss of timely data. Over the last 10 years, the Asynchronous Transfer Mode (ATM) networking has been progressively introduced. Early work on some non-critical systems established that ATM is reliable and flexible; therefore the network has now grown to include 31 systems and some 500 signals. The network itself has not lost a packet since its implementation and the problems have always taken place at the level of the single systems interfaced to the network. Fig. 3 shows the data connections from the Diagnostics to the Analysis (BetaLi etc.), to the Signal Server and Controller. Each system is connected to a network switch by a 155MB/s, two-way ATM link. Each switch has 16-ports, and a 2.5GB/s switching fabric. i.e. each switch can cope with all links running at full rate. The present JET network is implemented with three 16-port ATM switches. For further JET diagnostic Enhancements, 15 or more additional ports may be required. To cater for this expansion and to improve the integrity of the network, some additional ATM switches have been procured.

The JET real-time network infrastructure is advanced with respect to other fusion laboratories in the world. In the past it allowed the execution of very sophisticated real-time experimental programmes, particularly in the field of Advanced Tokamak (AT) scenarios, which require an integrated profile control in a multiple time scales. On the other hand, as the size of JET data is increasing, to transport large amount of data compatibly with JET lag time, high-speed data processing is required to meet the more demanding requirements and complex algorithms for plasma control, in time scales down to the sub-millisecond range in exceptional cases like the control of the vertical instabilities in high elongated plasmas. In addition, most of the new diagnostics being developed (neutron cameras, γ-ray cameras and hard X-ray cameras) will produce real-time data streams with rates up to 200MB/s/channel requiring an extension of the existing network infrastructure. Although the existing ATM technology used in JET RT network is sufficient to cope with these particular improvements, JET ongoing diagnostic upgrades constitute a singular opportunity to explore modern communication standards that are emerging in the computing and telecommunication industry. These standards take advantage of "switched-fabric" networks and "peer-to-peer" data interconnects and attain superior data rates and latency delay much superior comparing to ATM. Table 1 depicts current standard switch fabric based interconnects, all usable for a 10 μs control loop. Switched-fabric routing between acquisition/generator endpoints and data processors allows easy scalability of the number of channels while maintaining performance. This type of interconnection provides also important features like multicast/unicast delivery, priority based routing, support for tunneling of other standards (e.g. ATM) or proprietary protocols, among others. Such enhancements would extend significantly the real-time control capabilities of JET allowing faster distribution of measured and calculated plasma variables, higher complexity control algorithms processing input from multiple diagnostics and systems. In addition, it would allow the distribution of timing and synchronization over the real-time network and compatibility with existing networks, and to fully exploit the capabilities of modern hardware architectures such as Advanced Telecommunications Computer Architecture (ATCA), that is appearing now on JET in various applications (e.g., Neutron and γ-ray diagnostics, new systems for the control of the vertical plasma position, hard X-ray diagnostics etc.). This is particularly relevant

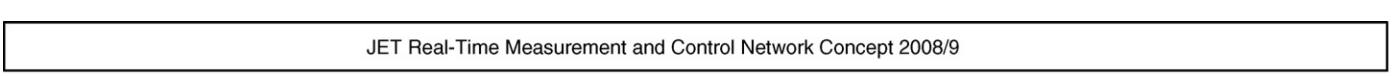

Fig. 3. JET real-time network based on ATM.

Table 1
Low latency switch fabric interconnects.

	Infiniband	ASI (PCIe)	Serial RapidIO	10GbE
Throughput (real max.)	197MB/s (1×) 790MB/s (4×) 2.37GB/s (12×)	205MB/s (1×) 820MB/s (4×) 2.46GB/s (12×)	275MB/s (1×) 1.1GB/s (4×)	763MB/s (4×)
Latency - switch - ping-pong	0.2 μs 2.4 (typ.)	0.1–0.5 μs 1.5 μs (typ.)	0.4 μs 1.2 μs (typ.)	<2.1 μs
Max length - copper - MMF	17 m (copper) 300 m	7 m (copper) 100 m	1 m (backplane) 100 m	15 m (copper) 26, 66, 300 m
Multicast, tunneling, Virtual Channels	Y	Y	Y	Y
Target application	Computer clusters	Networking	DSP farms	LAN, WAN
Base vendors	Mellanox	Stargen	Mercury, Tundra	Various
Availability	Now	2007	Now	Now
Cost	High	Moderate	Moderate	Moderate (dropping)

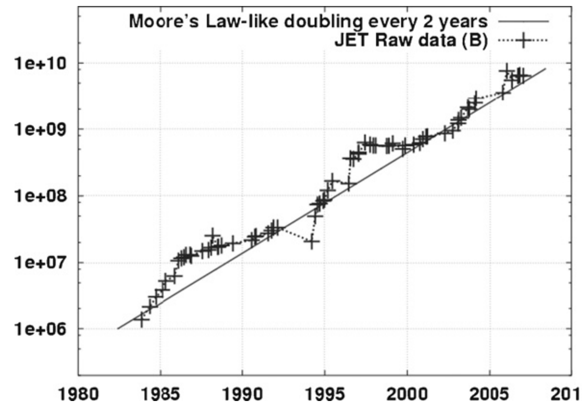

Fig. 4. The amount of data at JET follows Moore's Law.

for future fusion experiments with burning plasma which require stricter control of neutron rate and alpha particles with limited actuators. Such an infrastructure will constitute an ideal platform to test new concepts for ITER. To implement this upgrade, a preliminary study is under way to determine the most appropriate technology.

3. Data acquisition

From next campaigns at JET onwards towards the operation of ITER, the performance of the data acquisition systems will have to improve significantly. First of all, as mentioned already in Section 1, the shear volume of data to be acquired in present day machines is increasing at a rate exceeding the Moore law. This is shown in Fig. 4, which reports how the trend of acquired data per shot at JET has almost always presented a derivative exceeding doubling every two years. This trend does not show any tendency to decline and on the contrary, given the higher diffusion of video signals, it is expected to accelerate significantly. In order to cope with such an amount of data, compression techniques have to be refined. Since in a scientific environment, like fusion devices, the available information is in general very precious for the understanding of the phenomena under study, lossless techniques have to be developed. In general fusion machines, the signal information consist of a variable data header i.e., signal name, calibration factors, gain, etc., and a homogeneous part, which contains the digital codes of the samples. In some machines, both parts are initially stored without compression and the entire file is compacted afterwards using a suitable utility of the computer operating system. In contrast, in other devices data compression is only implemented over the homogeneous part of the signal information. This is the case of the delta compression techniques described in this paper and already implemented at JET.

The main principle of these methods is to store the differences between digital codes of adjacent signal samples instead of storing the digital codes of the original signals themselves. This approach, originally developed by Fredian at Massachusetts Institute of Technology [4], belongs to the family known as delayed methods, whereby an optimum bit allocation is found after examining the data. More recently, various refinements have been implemented, like the three-step algorithm used at JET and TJ-II [5,6]. The first step in this technique consists of a delta calculation, the second is a variable-length encoding of the resultant deltas before the resulting codes are stored in tandem. Since it has been determined, on the basis of a statistical analysis of empirical signals, that a typical delta distribution model can be assumed for each class of signals, a common encoding scheme can be adopted. Therefore this class of algorithms presents a signal independent bit allocation, which allows avoiding examination of the data as required for optimum

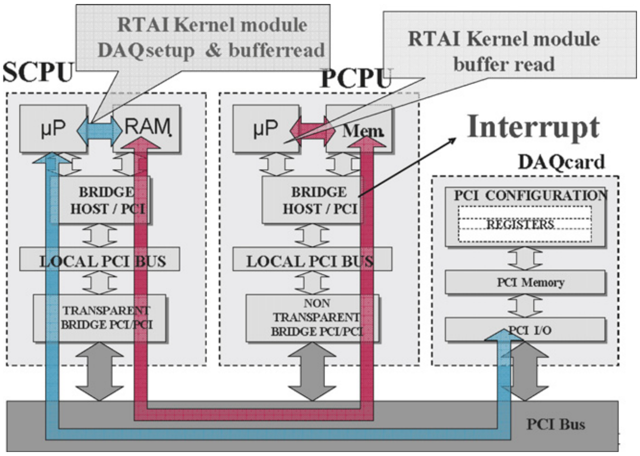

Fig. 5. Low level data flow.

bit length determination, resulting in an important reduction in processing time. Furthermore, this technique ensures the recovery of the initial signal without distortion when compacted data are expanded.

To reduce the data storage, other approaches are also being investigated. One alternative is based on event oriented measurements. In the version being implemented at JET, this approach consists of acquisition systems with an adaptive frequency. In this prototypical scheme, the data are sampled at a certain basic frequency and a first analysis is performed in real-time. If particular events are detected in the acquired signals, the acquisition frequency can be modified, typically increased, to better match the physical phenomena to be studied. In this way a significant reduction in the amount of data to be acquired can be achieved.

In the present JET architecture, an Intelligent Test and Measurement System (ITMS) [7] is implemented, which is based on the PCI eXtension for Instrumentation (PXI) standard platform, using DAQ cards, system CPU controller and peripheral CPUs. The architecture, oriented to fast data processing, was developed using Linux (RTAI), C language and LabVIEW. In more detail, the data acquisition was implemented with Linux (RTAI) using COMEDI (kcomedi), whereas the data processing is based on LabVIEW. The low level data flow is reported in Fig. 5. The data acquisition configuration (LabVIEW application) can be set up from either a user interface or using the extensible Markup Language (XML) language for remote management. The data acquisition and distribution (RTAI C kernel module) supports different cards. Both internal and external clocks and triggers can be used. The system guarantees continuous reading and the interface with RTAI FIFOS is implemented with LabVIEW CINs (Code Interface Nodes -C Language-).

The algorithm running on the hardware previously described is shown in Fig. 6. The sample rate clock is generated with the PXI system and the signals are collected in blocks. An asynchronous detection of sample rate change is implemented with the dynamic data processing system (DDPS) application. The decision to change the sampling rate can be defined by a user algorithm, on the basis of user-defined events, which can range from simple amplitude changes to more sophisticated phenomenology requiring recognition algorithms to be identified. The data reduction mechanism has been implemented using the ITMS-DDPS.

There are three types of frames to send data to network server(s):

(1) NFW: Normal float waveform. Waveform data in float format (SGL 32-bits) plus sample rate value.
(2) NBW: Normal binary waveform. Acquired Waveform in binary (up to 32-bits) format plus sample rate value.

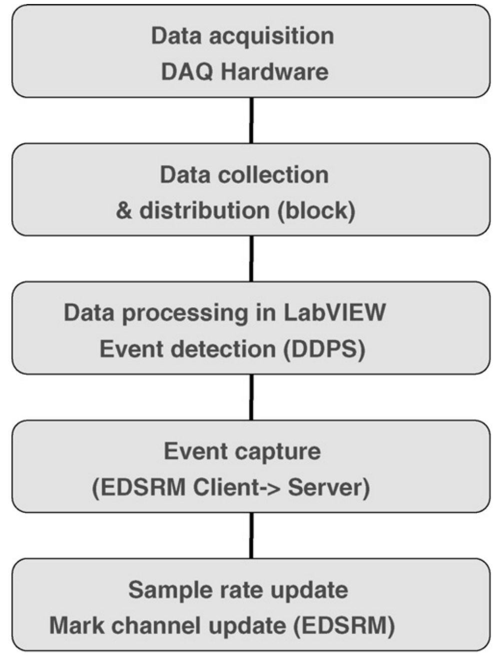

Fig. 6. ITMS software loop.

(3) CBW: Compressed binary waveform. The frame includes the sample rate value, the compression method used, the initial sample of the frame and the data (deltas).

Peripheral CPUs and System CPUs can send the compressed (or processed) information continuously in short length blocks to specialized computer systems. To further develop this approach for JET and to transfer it to ITER, several implications will have to be addressed. First of all, methodologies for event description, general enough to be portable to various diagnostics, will have to be developed. Then robust detection identification schemes will be necessary. From a more technical point of view, for these applications to become routine, the real-time networks will have to be able to handle the distribution of the events to all the interested parties with low latency and high reliability.

Another application in which intelligent storage of information has proved to be very important in the last years is the field of neutron spectrometry. In order to reduce the error bars in the determination of the measured spectra, acquisition rates must be high. On the other hand the information is relevant only in coincidence with the pulses corresponding to the arrival of neutrons. It is of no use to store the samples between successive events where there is absolutely no information in the signal. The storage space saved can be very significant without any loss of useful information. The prototype of a system implementing this approach has already been tested at JET [8]. The 14-bit 200 MSamples/s digital system consists of a high-speed PCI board (National Instrument 6534) and two 14-bit 100 MSamples/s ADCs (Analog Devices). The two ADCs are coupled in interleaved mode (in order to obtain an actual sampling rate of 200 MSamples/s). An Altera 1S25 field programmable gated array (FPGA) and a LabVIEWTM software devoted to acquisition handling and pulse analysis complete the architecture. In this application the analogue signal, typically the output of a photomultiplier, is digitized in discrete sample groups of variable-length (windows). A pre-programmed number of samples is acquired when the signal crosses a pre-set hardware threshold: however, the length of the time window is automatically increased to account for longer pulses or in case of further events occurring. This Dynamic Windows Data Acquisition logic approach allows acquiring only the data that actually carry information leading to a strong data compression that is of great help for the successive data transfer/storage.

4. Information retrieval and general analysis methods

In fusion devices, as in many other fields of science, the amount of available data is so large that nowadays the retrieval of the relevant information is a bigger issue than the collection of data. In JET, this has become particularly evident, since scientists from all over Europe participate in the experiments and the interpretation of their results. Already the preliminary task of identifying the most relevant discharges, to concentrate on them the subsequent analysis efforts, has become a very time consuming and sometimes frustrating work. In order to alleviate this problem and render the first data screening a much more efficient process, completely new way to organise the data bases is under intensive investigation [9]. In the new paradigm which is under test at JET and TJ-II, signals are not organized according to the traditional scheme of signal name, shot number and time slice. On the contrary, the objective consists of giving the scientists the opportunity to select, on a simple visual interface, the signal or a part of a signal they are interested in and then it is the task of the software system to extract the shot numbers and time intervals in which the same pattern is present. This new approach, which is meant to complement not to substitute the old database structure, is illustrated in Fig. 7.

The methods to achieve such a change in the underlying philosophy of database structure are derived from pattern recognition, which is a mathematical formulation of the more general problem of object identification. The recognition and classification of objects are characteristic goals of all living beings, since they are essential in basic activities like search for food or navigation, identification of friends and enemies etc. Pattern recognition identifies the group of disciplines which study the mathematical aspects of these activities. Since in plasma physics many phenomena can indeed be identified by the form of the corresponding diagnostic signals, in fusion the most promising techniques for the first data screening are the ones of structural pattern recognition. They are based on the general structural shape of the signals more than on their detailed mathematical properties (unlike correlation or coherence estimators), given the fact that these shapes in the signals are indeed what the physicists normally use to identify the various phenomena of interest during the visual inspection of the measurements, a preliminary step to the subsequent data analysis. The strategy followed at the moment therefore consists of developing methodologies that can perform a first screening of the data extracting the signal shapes which are similar to a prototype defined by the user [10]. The first step in achieving this goal is the extraction of the relevant features from the signals. On the basis of these features, a suitable distance has to be defined, which would allow assessing automatically the degree of similarity between different signals.

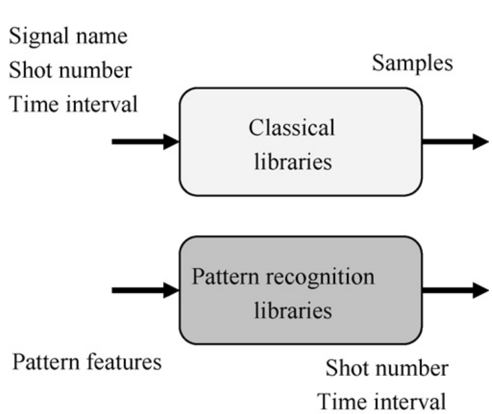

Fig. 7. Comparison between the traditional structure of fusion databases (top) and the new approach (bottom).

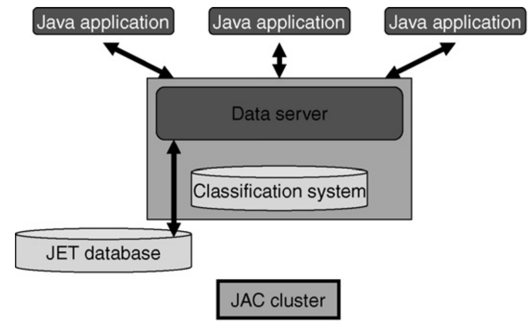

Fig. 8. Structure of the implementation at JET.

To exemplify these concepts the approach adopted to identify entire waveforms is particularly interesting [11]. A Haar wavelet transform is first applied to the signals in order to reduce of two orders of magnitude the dimensionality of the problem. On the basis of the features in the space of the transforms, the similarity measure used is the normalized dot product. In addition to admitting an easy geometrical interpretation, this choice of "distance" is particularly useful because it is independent of the signal amplitude and polarity. To increase the speed of the search, the various signals are classified in different groups depending on the length of the entire discharge. Therefore, for each signal to be retrieved, it is not necessary to traverse the whole database but only the subtree of discharges with the same length. A public version of the algorithm has already been implemented at JET (see Fig. 8). The results are very positive, since hundreds of signals can be retrieved on a subsecond time scale as illustrated in Fig. 9.

In Magnetic Fusion, the interpretation of the measurements is affected by various important uncertainties. First of all the plasmas depend nonlinearly on many parameters, some of which are not completely under control. Secondly, since plasmas are by their nature very difficult to access, many important internal quantities must be derived indirectly by the spontaneous plasma emissions (radiation and particles). These factors lead to quite complex inversion methods, in which the many measurements have to be considered to obtain the required information.

One recent development in the field of data analysis is the combination of the measurements from various diagnostics to obtain an unified identification of the plasma state. In the traditional approach, individual diagnostics are treated separately and, on the basis of their individual interpretation codes, they provide the value of a certain quantity often independently from the other measurements. On the other hand, using Bayesian statistical methods, the various instruments can be integrated in a single model and all contribute at the same time to the inference of the plasma state [12]. This holistic integration of various, and in principle all, the measurements in an unique model of the plasma constitutes the so called Integrated Diagnostic Modelling (IDM) approach which is being pursued now for the first time in a coherent and systematic way [12]. The extension of these methods would be also extremely beneficial to handle the level of uncertainties that always are inherent in the interpretation of new large experiments at the frontier of knowledge. This activity requires quite massive efforts since the forward functions of the diagnostics, which map the physics parameters to the expected measurements, must be evaluated. These forward functions have to include on the one hand all the calculations connecting the physics of the plasma to the final measurement and, on the other hand, all the sources of

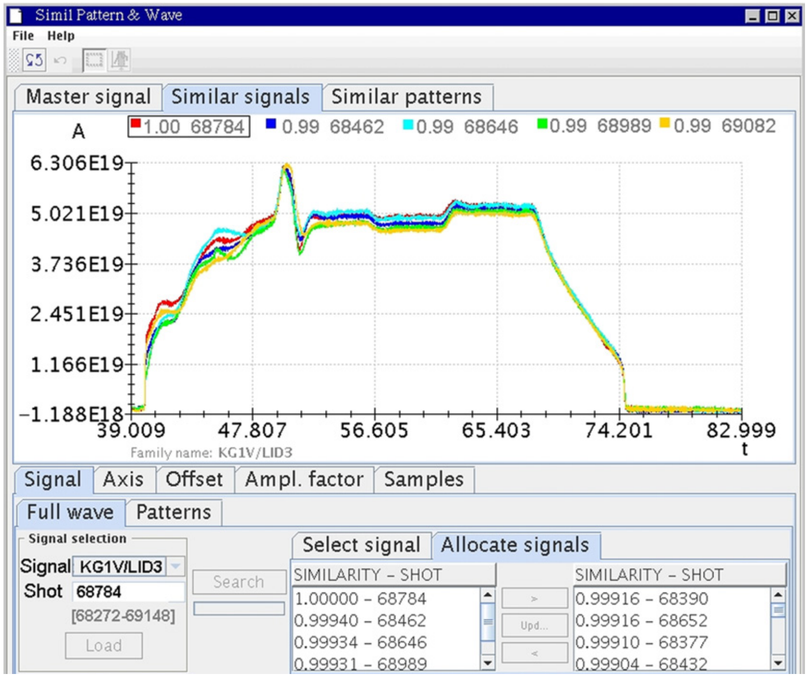

Fig. 9. Example of complete waveform retrieval for the case of JET plasma density as measured by the interferometer. Each query retrieves not a single signal but the set of the most similar waveforms to the initial one.

uncertainty of the diagnostic systems: calibration factors, positioning parameters etc. Since the most complex diagnostics can have dependencies on hundreds of parameters, a general architecture is necessary to properly handle this amount of information. Two architectural infrastructures have been put in place to implement such an IDM programme in JET. The first, called MINERVA, is a generic framework that formalises the specifications of diagnostic models, parametric dependencies, and their relationship to physics models, allowing optimal combination of measurements and optimal Bayesian inference on physics models. The other aspect of these integration activities consists of the large calculations which are sometimes necessary to relate the raw measurements to the physical model, like for example the tomographic inversions. These are often implemented in large legacy codes, typically written as stand-alone FORTRAN programs. The accessing, understanding and running these legacy codes is currently complicated and creates a barrier for the general user, not expert of a specific code. The general programme of IDM has therefore to face the problem of combining such codes with each other, for example the output of one code being used as input or partial input to other codes. At present this could easily require months of work for just a couple of programs. The SOFI (Service Oriented Fusion Initiative) project is trying to remedy this unsatisfactory situation by writing codes according to a so called 'Service Oriented' Architecture (SOA), using Web Service technologies. To achieve the compatibility between codes, all data communicated between codes is strictly formatted in language-independent XML Schema (XSD) structures and the functional interface between codes and between users and codes are defined using language-independent XML WSDL (Web Service Description Language) documents. With this approach the interaction between codes can be defined in a language-independent and machine-readable way, which among other things makes much easier their parallelisation.

5. Remote participation and remote experimentation

The concept of remote participation involves the development of highly interactive, powerful, efficient and secure shared environments with the aim of achieving the following goals:

(1) Real-time participation
  - Meetings: collaborative tools for personal interactions
  - Device operation (long pulse)
    - o Interplay with technical/experimental systems
    - o Monitoring/programming purposes

(2) Off-line data access
  - Data analysis, integration and retrieval

At this point, it is worth mentioning that 'remote experimentation' is a more restricted concept than remote participation. From our point of view, remote participation includes all the above points but remote experimentation is only related to 'Device operation'. Following the operation under long pulse conditions demands a strong interplay with control, data acquisition and diagnostic systems in order to obtain the relevant information at any moment and to be able to modify on the fly pre-programmed configurations. For example, it could be required to change a sampling rate, to stop signal digitization or to start data sampling in some channels.

Remote experimentation means to be able to perform *on-line* interactions with the experimental systems to 'remotely' *follow* the temporal evolution of magnitudes and *command* configurable parameters. It should be highlighted that the term 'remotely' may refer to both the control room local area network (LAN) and a remote site LAN.

A remote experimentation model has to be centred on providing tools for firstly, real-time interactions with diagnostics, data acquisition and control systems and secondly, real-time data visualization. The design of a remote experimentation model has to ensure the creation of an operation environment with the following characteristics: multi-platform, scalable, high security and effortless administration. Web technologies offer a good framework for the development of a remote experimentation system due to their open character, technological maturity, standardization level, security properties, scalability system architecture, and platform independency.

5.1. Design principles of a remote experimentation model

Several main design principles have to be taken into consideration according to the characteristics of both the experimental environment and the user needs. One of these is the remote experimentation requirements and the firewall security issue. The authentication/authorization application consistent with the presence of strong firewalls is a very important issue with regard to the success of remote experimentation technologies.

Another principle is to ensure that local and remote participants could use the same general tools for system monitoring, system programming and visual data analysis. Two consequences are derived from this principle. First, the model treats all user the same. Second, execution of multiple developments for the same purposes is avoided, thereby ensuring easier maintenance.

An additional point is to forbid the use of fat clients for remote experimentation. Contrary to thin clients that provides user interface only, a fat client is a piece of software that runs on a client computer and includes both code to process data and code to present it. This can lead to security and software distribution problems. Concerning security, fat clients contain 'business logic'. This business logic may include code that accesses tables and fields directly in a business database and it should be avoided. Regarding software distribution, any change in the software compels to update every time all client software in order to preserve data integrity.

Last, but not least, requirements for the security system are mentioned. It is essential the use of a distributed authentication and authorization system as a consequence of the distributed nature of the remote experimentation system. It is necessary to deal with client software worldwide spread and it is also necessary to manage the access to different experimental systems within the local area network environment.

5.2. Multi-tier architecture for remote experimentation applications

Our remote experimentation model establishes thin clients as the user applications. It means the generation of a multi-tier software architecture based on three layers. One tier (client tier) is devoted to providing user interface code only (this is the thin client). The second tier (middle tier) provides resources for authentication, authorization and query executions. Finally, the third tier (data tier) is responsible for the management of the data to be used by the client tier. A multi-tier architecture like this tries to prevent build-time dependencies between tiers so that one can update the software on one tier without having to update the software on its neighbouring tiers, which can be essential for remote applications.

It is important to understand the implications of this architecture in collaborative applications for the interplay with data acquisition systems and diagnostic control systems. Collaborative applications are applications that work in a cooperative manner and share data. An example of this can be a data acquisition system (Fig. 10). First, we have the data acquisition system itself with its control software for data taking. Second, we have a client application which can modify acquisition parameters. Third, we have other client application with read only access to the system. All applications follow the three-layer model and they work in a cooperative way. Communications *only take place by means of the data tiers* that reside necessarily within the experimental environment. Moreover, the authentication and authorization system in the "Middle tier" could be different between applications. For example, the control software is executed in the experimental LAN and its security system could be different from the one concerning client applications.

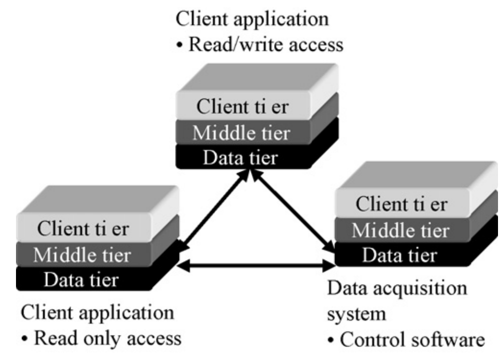

Fig. 10. Collaborative applications communicate only through the data tier.

5.3. Data tier, middle tier and client tier

The data tier of an application is fundamentally a data server that contains exclusively relevant data to its client tier. Collaborative applications share data through the data tiers. We can imagine the several data tiers of collaborative applications as views of a relational database. It focuses the important data for the application and ignores unnecessary information. The data tier software always must be in execution in the local area network environment. Also, it should be underlined that the data tier can manage its own security method for data access, independent of the authentication and authorization mechanism of the "Middle tier". For instance, we can think of a database that has defined its own usernames for access.

The "Middle tier" is the interface layer existing between users and data. It has to be integrated together with the general authentication and authorization system to grant or not the access to the "Data tier". This layer also translates the user queries in such a way that the "Data tier" can understand. In addition, the "Middle tier" has to provide "connectors" to link the general authentication and authorization system to the "Data tier" security methods (if any).

The client tier is really a thin client application because it is only a presentation layer. Client tier software can be Web pages or any kind of application (C/C++, 4GLs or Java). The existence of a client tier allows the development of different views for the same data. Moreover, it makes software deployment easier.

Software distribution and version control can be a problem in local area networks, which is amplified outside the local environment. Thus, it would be necessary to establish automated software deployment methods to ensure the execution of the last version of the applications. For example, the Java Web Start software, which is a standard component of the Java Runtime environment, can be used for this.

5.4. Discussion on thin client environments

Thin clients and automated software download are a very powerful combination to make remote experimentation easier. In fact, it ensures the execution of the last version of the different applications without increasing the system administration overhead because the application download is carried out at the execution time. Furthermore, thin clients can be applications as complex as needed to integrate as much functionalities as possible, but, in general, the power of the thin client approach is to create several small pieces of software instead of having a single one managing the whole environment. If the thin clients are Java applications or web pages, then a real multiplatform environment is built and only a single development is needed, even to be executed from the LAN.

In addition, thin clients enable the creation of multiple views for the same data. The existence of thin clients also permits to achieve the highest degree of security. Security rules are not coded inside user applications but they reside in the protected experimental environment. Moreover, the multi-tier architecture filters the access to only the required data of each application, thus preventing the sharing of unnecessary data.

Other important fact is that the computation power is distributed. Application programs are executed in client computers. Also, it should be remarked that the thin client approach is optimal for real-time systems. We incur minimum overload on experimental systems since the systems themselves do not have to provide access control, database resources or graphic user interface capabilities.

Finally, it should be mentioned that the TJ-II remote participation system was based on thin clients [13]. The development has been used on a routine basis (even from the TJ-II local area network in the daily operation) since five years ago.

6. Conclusions and future developments

In future generations of fusion devices, the requirements in terms of machine integrity and long pulses will impose a significant change of paradigm in the control of the experiments. Diagnostics will have to be more reliable with sound in situ periodic recalibration techniques to guarantee the quality of the measurements. The real-time control system will be systematically in charge of running the experiment and much less recourse to feed-forward schemes will be allowed. To this end, significant optimisation of the data to be distributed will have to be performed. With regard to the offline analysis, the enormous amount of information acquired per discharge will have to be analysed in more efficient and automatic ways than at present, to provide timely inputs to the planning of the following discharges. [14]

In this context, in addition to the technologies and methods described in this paper, some other innovative approaches are planned to be tested at JET. Since some of the most ambitious real-time control schemes have been proved to be very heavy in terms of machine time, Fuzzy Logic could be an alternative to speed up the definition and training of controllers. This would a pioneering application in Nuclear Fusion but unfortunately it has not been undertaken so far. Given the fact that cameras are used more routinely than in the past, more sophisticated methods to process images will have to be developed and tested. They will include not only automatic image processing algorithms but also more sophisticated methods to interpret the information. Automatic updates of the regions of interests, for infrared thermographic surveys, are just an example of applications requiring significant developments compared to present day devices. The integration of various measurements in the decision loops is also an aspect which will have to be addressed more systematically.

Acknowledgements

This work was partially funded by the Spanish Ministry of Science and Innovation under the Project No. ENE2008-02894/FTN.

This work, supported by the European Communities under the contract of Association between EURATOM/CIEMAT, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] R. Neu, R. Dux, A. Kallenbach, T. Pütterich, M. Balden, J.C. Fuchs, A. Herrmann, C.F. Maggi, M. O'Mullane, R. Pugno, I. Radivojevic, V. Rohde, A.C.C. Sips, W. Suttrop, A. Whiteford, The ASDEX Upgrade team, Nuclear Fusion 45 (3) (2005) 209–218.
[2] J. Paméla, G.F. Matthews, V. Philipps, R. Kamendje, JET-EFDA Contributors, Journal of Nuclear Materials 363–365 (2007) 1–11.
[3] R. Felton, E. Joffrin, A. Murari, Fusion Engineering and Design 74 (1) (2005) 769–774.
[4] T.W. Fredian, J.A. Stillerman, Review Scientific Instruments 57 (8) (1986) 1907–1909.
[5] J. Vega, C. Crémy, E. Sánchez, A. Portas, S. Dormido, Review of Scientific Instruments 67 (12) (1996) 4154–4160.
[6] J. Vega, M. Ruiz, E. Sánchez, A. Pereira, A. Portas, E. Barrera, Fusion Engineering and Design 82 (2007) 1301–1307.
[7] M. Ruiz, J.M. López, G. de Arcas, E. Barrera, R. Melendez, J. Vega, Fusion Engineering and Design 83 (2008) 358–362.
[8] D. Marocco, M. Riva, B. Esposito, A. Zimbal, International Workshop on Fast Neutron Detectors and Applications (Cape Town, April 2006), PoS (FNDA2006)028, http://pos.sissa.it/.
[9] J. Vega, JET EFDA Contributors, Fusion Engineering and Design 83 (2008) 382–386.
[10] J. Vega, G.A. Rattá, A. Murari, P. Castro, S. Dormido-Canto, R. Dormido, G. Farias, A. Pereira, A. Portas, E. De la Luna, I. Pastor, J. Sánchez, N. Duro, R. Castro, M. Santos, H. Vargas, Proceedings of the IEEE International Symposium on Intelligent Signal Processing, 2007, pp. 949–954, ISBN: 1-4244-0829-6.
[11] J. Vega, A. Pereira, A. Portas, S. Dormido-Canto, G. Farias, R. Dormido, J. Sánchez, N. Duro, M. Santos, E. Sánchez, G. Pajares, Fusion Engineering and Design 83 (2008) 132–139.
[12] J. Svensson, A. Werner, Proceedings of the IEEE International Symposium on Intelligent Signal Processing, 2007, pp. 955–960, ISBN: 1-4244-0829-6.
[13] J. Vega, E. Sánchez, A. Portas, A. Pereira, A. Mollinedo, J.A. Muñoz, M. Ruiz, E. Barrera, S. López, D. Machón, R. Castro, D. López, Fusion Engineering and Design 81 (2006) 2045–2050.
[14] J. Vega, E. Sánchez, A. Portas, A. Pereira, A. López, E. Ascasíbar, S. Balme, Y. Buravand, P. Lebourg, J.M. Theis, N. Utzel, M. Ruiz, E. Barrera, S. López, D. Machón, R. Castro, D. López, A. Mollinedo, J.A. Muñoz, Fusion Science and Technology 50 (2006) 464–471.