Common browsing and retrieval of technical and experimental data for steady state experiments

Torsten Bluhm[a,*], Peter Heimann[b], Christine Hennig[a],
Georg Kühner[a], Hugo Kroiss[b], Heike Laqua[a], Marc Lewerentz[a],
Josef Maier[b], Heike Riemann[a], Jörg Schacht[a], Anett Spring[a],
Andreas Werner[a], Manfred Zilker[b]

14th July 2009

[a] Max-Planck-Institut für Plasmaphysik, EURATOM Association, Teilinstitut Greifswald, Wendelsteinstraße 1, D-17491 Greifswald, Germany

[b] Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstraße 2, D-85748 Garching, Germany

[*] Corresponding author; Tel.: +49-3834-88-2238 E-mail address: Torsten.Bluhm@ipp.mpg.de

Abstract

Steady state experiments like Wendelstein 7-X have to deal with large amounts of data of different kind. Discharge lengths may last from less than a second up to 30 minutes. During this time diagnostics will measure data in different discharge phases with different rates depending on the physical program. Furthermore the technical configuration of all devices must be logged during the discharge as well as between experiment phases. This is necessary to allow supervision of constantly running devices. All measured data must be stored together in a common database to make it easily accessible and reduce the effort to combine data. Because of these requirements a continuous data acquisition system must be provided where data may exist at any time. Intelligent methods for browsing and retrieval of the acquired data from the common database are desired. This includes calendar based time interval selection, overview plots for intervals that cover several days and online views of the last stored data. The contribution will explain the challenges and the advantages induced by this kind of data access methods. Existing and planned solutions for Wendelstein 7-X will be depicted. Particularly the "Data Browser" - a graphical tool specially designed to fulfil these tasks at W7-X - will be presented.

Keywords: Wendelstein 7-X, continuous operation, data access

1   Introduction

Data retrieval at large scientific experiments like Wendelstein 7-X is a complex task. Like on any other experiment this topic has to be handled in as efficient as possible, as correctly as possible and - not to forget - as user friendly a way as possible. However, there are some differences compared to other fusion devices that are related to the steady state operation of the machine combined with continuous data acquisition. Experiment durations of up to 30 minutes will open up exciting possibilities to observe and manipulate the plasma even *while* the machine is running. Therefore monitor and analysis tools must take into account new challenges, for example arbitrary parameter changes during an experiment and improved data management as a consequence of the huge data blocks that have to be processed.

A basic principle at W7-X is that all data obtained from the experiment is saved in a common database. This includes

- measurements of physical quantities like temperatures, intensities or phase shifts
- measurements of technical quantities like currents of power supplies

- parameters that are used to control the hardware and software components

- machine and component states and

- experiment process information.

All this information is necessary to get a coherent view of the experiment and to be able to do a meaningful data analysis. Given that fact, the crucial question is how this data mass can be presented to the user in a comprehensible and manageable way.

2 Experiment views

2.1 General ideas

People working at W7-X are specialists in their domain and therefore have different ideas of what data is to be displayed and how it should be represented. Therefore we have to consider domain-specific views which additionally depend on the particular environment.

Generally we have three different "roles" that will be involved in the experiment:

1. the experiment leader

2. the diagnostician and

3. the technical operator

There may be other roles marginally interested in the experiment like remote participants but those will not be considered here.

For all three of the above roles we have to distinguish between two use cases:

1. Online - which means the span of time while an experiment is running and

2. Offline - which describes the time after an experiment has finished.

For the Online case the experiment leader will be interested in things like the experiment process, the overall machine state and maybe some leading experiment quantities. The diagnostician on the other hand will want to know the individual state of his particular diagnostic and some characteristic quantities. He may also want to do some data analysis in the form of feedback algorithms or routine analysis. While the outputs of feedback algorithms directly influence the plasma the results from routine analysis permit efficient ways to access data in subsequent analysis processes. The technical operator finally will need a more detailed view of the component in his responsibility including the detailed state and safety signals.

Offline, the experiment leader needs to be able to review previous experiments. To get a quick impression of an experiment some kind of overview plot would be appropriate that shows simple time traces of the leading experiment quantities. He will also need suitable search functions to find similar experiments given some configured, measured or analysed quantity. This may involve complex pattern searching algorithms as they are currently under investigation[1, 2]. The diagnostician will need all this as well and additionally require methods to get a more detailed view of the data, to access data with the analysis software of his choice and to produce publishable plots based on these analyses. The technical operator will most likely focus on comparisons between set-points and actual values in order to fine-tune the hardware for the next experiment.

2.2 Realisations at Wendelstein 7-X

All the requirements described in the previous chapter are meant to be covered by the integrated control and data acquisition system of W7-X[3, 4]. Figure 1 illustrates the structural overview of the system and its interfaces to the described roles.

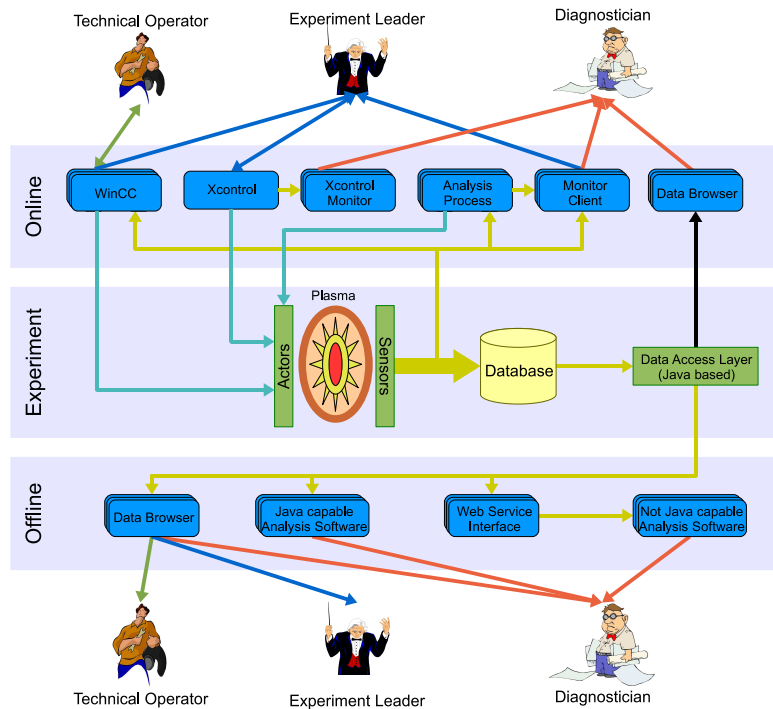The major applications that will be used for online observation are:

Figure 1: Experiment views

- Xcontrol - the user interface for the W7-X segment sequence control. Besides the control instance that is used by the experiment leader to define the experiment sequence several monitor instances can be opened that mirror the experiment process for passive participants[5, 6].

- WinCC - the standard user interface for Programmable Logic Controllers (PLC). These applications display the state of the operational management and provide control methods for elements not covered by the segment control system. This interface will generally be used by the technical operators[7, 8, 9].

- Monitor Client - a client application for the W7-X monitoring system. These applications can present special pre-processed data streams to the user that are distributed over the network. Because band width is limited these are normally reduced versions of the original data streams[10].

As soon as data is transferred to the database it can be loaded and visualised with the *Data Browser* application. Typically data is transferred to the database every few seconds. For the case that the experiment lasts several minutes this means that data can be displayed with the *Data Browser* while the data acquisition is still running.

However, the main area of application for the *Data Browser* is the offline case. Because of its importance for offline data browsing this application will be described more detailed in the following chapter. For users who want to import data into their own applications it is of course also possible to integrate the *Data Access Layer* interface[11] which is used by the *Data Browser*. Since this is a Java based interface it can only be used by Java capable applications. For that reason it is planned to provide the same interface as a web service which makes it available to nearly every programming language. On the other hand the web service approach also allows the integration of further data repositories. A service for the new continuous MDS+ interface[12] could be developed as well as a service for shot based archives like the one used at W7-AS for example.

3 Navigating large data sets with *Data Browser*

The *Data Browser* application is meant to be a helpful tool to navigate both data acquired with a high resolution and data acquired over a long time interval. It is based on the Eclipse framework and therefore can be developed in a standardised and platform independent way. Figure 2 shows a typical view of a *Data Browser* window. The main window consists of several sub windows that can be arranged freely by the user. There are two types of sub windows: selections and plots. A selection is used to specify what data is to be presented while a plot actually presents the data to the user.
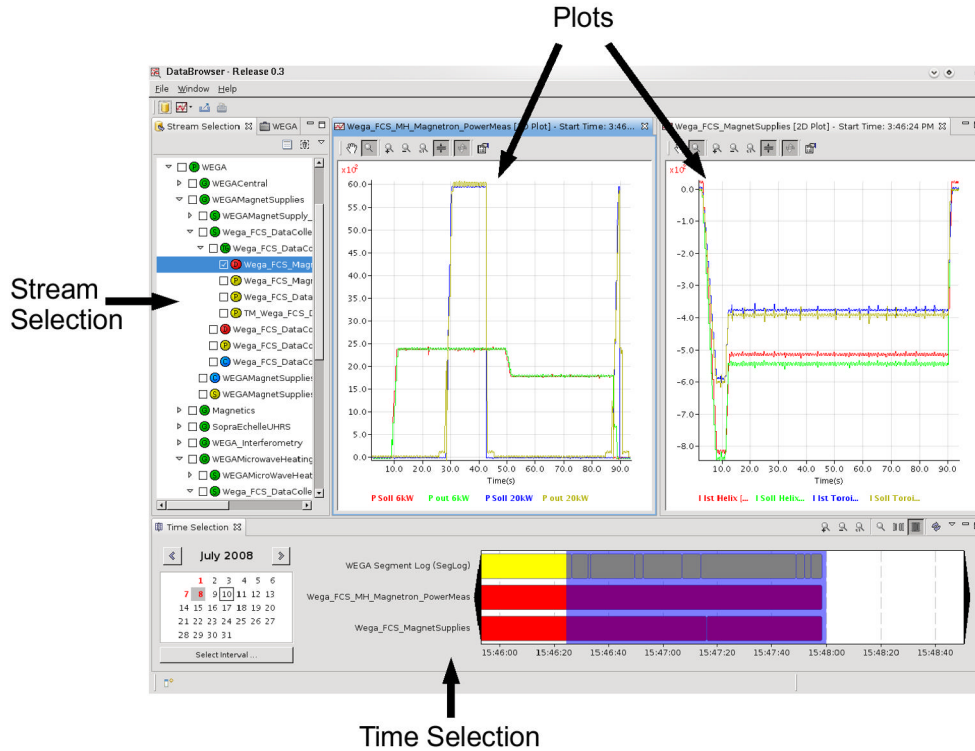

Figure 2: Data Browser screenshot

3.1 Data selection

The task of selecting data can be divided in two sub tasks:

- what data are we interested in and

- in which time interval are we interested

The two tasks are orthogonal to each other and therefore are represented by two separate selections inside the *Data Browser* main window.

The *Stream Selection* displays the experiment structure corresponding to the W7-X function group concept[13]. The complete function group hierarchy including the structure of its sub-elements (control and data acquisition stations) is represented in form of a tree below the database root node similar to the directory structure of a file system. The leaf nodes of this tree provide access to the acquired data. They can be marked to select the referenced data stream for further processing in the *Time Selection* which will be explained below. Of course such an experiment structure is not static and can change over time. For this reason it is possible to define the point in time for which the experiment structure is to be displayed. By default the last valid structure will be shown.

4

The *Time Selection* displays information about when data is available and provides possibilities to select a specific time interval. Which data streams will be taken into account depends on the marked nodes in the *Stream Selection*. Because we are using a continuous data acquisition system there are no fixed time intervals identified by a number like in shot based experiments. Instead we have to specify the time interval in that we are interested by defining the start and end of this interval. There are two elements in this selection: a calendar view and a Gantt-chart-like diagram view. In the calendar view all days for which data in the selected data streams exists are marked red. It is possible to select time spans that cover a single day or even several months for long term overviews. The selected time interval will be displayed in the diagram view where for each selected data stream the time intervals containing valid data will be shown as coloured regions along a time axis. Here it is possible to select the actual time interval that defines the data section to be visualised in plots.

Another selection that is in fact an addition to the *Stream Selection* is the Workspace. Workspaces are a common concept in software developed for W7-X and occur also in other applications. They can be compared to a favourites folder in a web browser. Data streams that are often used can be moved from the *Stream Selection* to a workspace to avoid navigation through the whole experiment structure. The items in a workspace work very similar to the tree nodes in the *Stream Selection*. The only difference is that they are arranged in a list rather than a tree. For every use case a new Workspace can be created. For example one could create a Workspace for experiment wide values and other Workspaces for component specific content.

3.2 Data plotting

Once data has been selected it can be plotted using one of several plot types. The plot type depends on the data type on the one hand and the user needs on the other hand. For example camera images have to be displayed differently than data acquired with an ADC. Furthermore the same data can be represented in different ways like ADC-acquired data can be visualised as a line graph as well as concrete values in a table. If the standard plots provided by the *Data Browser* are not sufficient it is possible to implement further plot types that can be added to the application as plug-ins.

The provided plots also support interaction with the data. For instance it is possible to zoom graphs and move graphs around to get a more detailed look at the data. Changes of the selected time interval in the *Time Selection* will be immediately reflected by an update of the displayed data. It is also possible to display more than one data stream in one plot and to select a subset of the channels in a data stream.

3.3 Data export

To make the acquired data available to other applications it may be necessary to export data to files in different formats. The simplest case that is already implemented is to store it in ASCII-files in a textual form. These files can be easily imported in standard spreadsheet applications for example. Like user defined plots other export methods can be implemented and integrated as plug-ins.

4 Conclusions and Outlook

The described data retrieval interfaces have already proven their usability in laboratory installations and test environments. A broader approach to test these applications is the W7-X prototype for control and data acquisition at the WEGA experiment[14, 15]. The principle applicability could be shown there and many details that still need improvement could be detected. This especially concerns the usability of the graphical user interfaces and performance issues.

The *Data Browser* is by now in a more or less stable state and fulfils the basic data browsing tasks as described above. To make it a robust and user friendly application a number of essential enhancements and additions have still to be implemented. The major ones are named here, providing a short overview.

1. A very important issue is the question of data searches. Based on given configured or measured values it must be possible to find the corresponding data in a larger data set. For example the search for equivalent parameter configurations in former experiments will be a crucial point. There are two approaches that address this task:

   (a) Implementations of database queries using an adequate query language
   (b) Using analysis algorithms to mark time intervals and browse the marked intervals with the *Data Browser* afterwards

2. To describe special time intervals inside a measurement more closely it would be useful to provide an annotation function. The user could select a data interval and enter a comment for this interval that would be stored together with the data in the central data storage. Such other interested users could also read these comments and optionally add their own ones. This corresponds to 1b except that the interval marks are created interactively rather than automatically.

3. The representation of the experiment structure in the *Stream Selection* is currently very strictly connected to the technical setup. It would be better to introduce a view that is related more closely to the experiment on a more abstract level. The ideal case would be switchable views that adapt to the user environment.

4. The way multiple plots can be arranged in one sub-window should be more flexible. This is especially useful when different measured values for a time section have to be examined together. It must also be possible to save these arrangements to provide the same layout for future experiments. Furthermore a print function for these kinds of plots would be useful.

5. The functionality of the Workspaces could be extended to become a tool for communication between users. This would require additional elements inside Workspaces like interesting time intervals, local notes or the like.

References

[1] J. Vega and JET-EFDA Contributors, Intelligent methods for data retrieval in fusion databases, Fusion Engineering and Design, Volume 83, Issues 2-3, Proceedings of the 6th IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation for Fusion Research, April 2008, Pages 382-386

[2] H. Nakanishi, T. Hochin, M. Kojima, LABCOM group, Search and retrieval method of similar plasma waveforms, Fusion Engineering and Design, Volume 71, Issues 1-4, 4th IAEA Technical Meeting on Control, Data Acquistion, and Remote Participation for Fusion Research, June 2004, Pages 189-193

[3] Jörg Schacht, Heike Laqua, Marc Lewerentz, Ina Müller, Steffen Pingel, Anett Spring, Andreas Wölk, Overview and status of the control system of WENDELSTEIN 7-X, Fusion Engineering and Design, Volume 82, Issues 5-14, Proceedings of the 24th Symposium on Fusion Technology - SOFT-24, October 2007, Pages 988-994

[4] P. Heimann, S. Heinzel, Ch. Hennig, H. Kühntopf, H. Kroiss, G. Kühner, J. Maier, J. Reetz, M. Zilker, Status report on the development of the data acquisition system of Wendelstein 7-X, Fusion Engineering and Design, Volume 71, Issues 1-4, 4th IAEA Technical Meeting on Control, Data Acquistion, and Remote Participation for Fusion Research, June 2004, Pages 219-224

[5] Anett Spring, Heike Laqua, Jörg Schacht, User control interface for W7-X plasma operation, Fusion Engineering and Design, Volume 82, Issues 5-14, Proceedings of the 24th Symposium on Fusion Technology - SOFT-24, October 2007, Pages 1002-1007

[6] Anett Spring, Heike Laqua, Helmut Niedermeyer, User interaction concept for plasma discharge control on WENDELSTEIN 7-X, Fusion Engineering and Design, Volume 81, Issues 15-17, 5th IAEA TM on Control, Data Acquisition, and Remote Participation for Fusion Research - 5th IAEA TM, July 2006, Pages 1957-1961

[7] Jörg Schacht, Helmut Niedermeyer, Heike Laqua, Anett Spring, Ina Müller, Steffen Pingel, Andreas Wölk, Tasks and structure of the WENDELSTEIN 7-X control system, Fusion Engineering and Design, Volume 81, Issues 15-17, 5th IAEA TM on Control, Data Acquisition, and Remote Participation for Fusion Research - 5th IAEA TM, July 2006, Pages 1799-1806

[8] PLC/SIMATIC systems by Siemens: http://www.siemens.de/simatic

[9] WinCC process visualization software by Siemens: http://www.siemens.de/wincc

[10] Ch. Hennig, P. Heimann, S. Heinzel, H. Kroiss, G. Kühner, H. Kühntopf, J. Maier, J. Reetz, M. Zilker, A concept of online monitoring for the Wendelstein 7-X experiment, Fusion Engineering and Design, Volume 71, Issues 1-4, 4th IAEA Technical Meeting on Control, Data Acquistion, and Remote Participation for Fusion Research, June 2004, Pages 107-110

[11] T. Bluhm, P. Heimann, Ch. Hennig, H. Kroiss, G. Kühner, J. Maier, H. Riemann, M. Zilker and EURATOM Association, Data access and its implementation at Wendelstein 7-X, Fusion Engineering and Design, Volume 83, Issues 2-3, Proceedings of the 6th IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation for Fusion Research, April 2008, Pages 387-392

[12] T. Fredian, J. Stillerman, G. Manduchi, MDSplus extensions for long pulse experiments, Fusion Engineering and Design, Volume 83, Issues 2-3, Proceedings of the 6th IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation for Fusion Research, April 2008, Pages 317-320

[13] Heike Laqua, Helmut Niedermeyer, Jörg Schacht, Anett Spring, Real-time software for the fusion experiment WENDELSTEIN 7-X, Fusion Engineering and Design, Volume 81, Issues 15-17, 5th IAEA TM on Control, Data Acquisition, and Remote Participation for Fusion Research, July 2006

[14] Jörg Schacht, Dieter Aßmus, Torsten Bluhm, Andreas Dinklage, Stefan Heinrich, Christine Hennig, Uwe Herbst, Ralf König, Heike Laqua, Marc Lewerentz, Ina Müller, Matthias Otte, Steffen Pingel, Jürgen Sachtleben, Anett Spring, Andreas Werner, Andreas Wölk, Stellarator WEGA as a test-bed for the WENDELSTEIN 7-X control system concepts, Fusion Engineering and Design, Volume 83, Issues 2-3, Proceedings of the 6th IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation for Fusion Research, April 2008, Pages 228-235

[15] Marc Lewerentz, Dieter Aßmus, Torsten Bluhm, Stefan Heinrich, Christine Hennig, Uwe Herbst, Christiane Meyer, Eric Köster, Ina Müller, Heike Laqua, Matthias Otte, Steffen Pingel, Jürgen Sachtleben, Jörg Schacht, Anett Spring, Andreas Wölk, First experiences with the new W7-X like control system at the WEGA Stellarator, This conference