Chapter 1

Control, data acquisition and retrieval in international experiments

1.1 JET

The Joint European Torus (JET) is currently the largest tokamak in operation in the world. It was designed in the mid 1970s and its operation began in 1983. It is installed near Culham, Oxfordshire in the UK and is managed through the European Fusion Development Agreement (EFDA). Not only was here that the first controlled release of fusion energy happened, but it also detains the world record in fusion power (16 MW). In 2006 operation was started with an ITER-like magnetic configuration and in 2011 an ITER-like wall was installed, thus becoming the key on the preparation for ITER operation. For the scope of this thesis, JET represents a good example of a large experiment that has been running for several decades, during a time of significant technological advancements.

At the begining, its CODAS was designed around a modular, tree-like structure. It consisted of three levels:

- Level 1: Jet-wide supervisory control
- Level 2: Subsystem supervisory control
- Level 3: Local unit or component

Level 2 and most of level 3 software ran on level 2 computers, with some real-time control on microprocessors in level 3. This structure was also used to identify data in the database. The technology used was cutting edge for the time, with NORD computers and CAMAC front-end electronics. Most software was written in Fortran, with time-critical software written in Assembly and NPL (NORD Programming Language). Around 1990, this technology was showing its age: more powerfull hardware and modern software were available, but they were not supported on the NORD architecture. It was thus decided to switch to a UNIX based system (Solaris OS), to use ethernet networking and to replace the CAMAC system with VME and PC front-ends. This was no easy task: there were many CAMAC systems being used and they could not be replaced overnight, so they had to be interfaced with the new computers; tens of thousands of lines of code written in Assembly and NPL had to be ported to C; around half a million lines of Fortran code also had to be translated because of NORD extensions that were not supported by UNIX; some NORD features also had to be temporarly emulated; and all of this while the machine had to remain operational. Only in 1994 and after a major shutdown phase, was the migration to the new CODAS system considered complete.

This, as expected, would not end here and by 2008 the system had received another overhaul, being by then mostly based on PCI eXtension for Instrumentation (PXI) and LabVIEW, with computers now using GNU/Linux operating systems, specifically Linux RTAI for real-time applications. More modern hardware architectures, like Advanced Telecommunications Computer Architecture (ATCA) here started to being used for new subsystems. In 2010, a real-time control framework, called MARTe (Multi-threaded Application Real-Time executor), that can take advantage of multicore processors, was developed at JET and now is being used for some of its real-time control, including some critical subsystems like the vertical stabilisation control.

It is also interesting to look at how the system grew during its opertional history. In the early design it was estimated that the CODAS would require about 15 computers, together with a large mainframe for data storage and analysis. It was also expected that this would be enough for the entire life of JET. But as soon as JET became operational this number had already grown to 34 NORD minicomputers (20 online), in 1999 there were 122 SPARC computers being used (92 online) and the number continued to grow to the point were it is not easy to make an estimation anymore. As for the data generated, it was first expected that each shot would produce almost 1 MB of data, value that was surpassed as soon as JET started regular operation. Over the years the amout of data collected has been steadily increasing, roughly doubling every two years, reaching 10 GB by 2008. In order to cope with this growth the first measure taken (besides improvments made on the network) was to classify data has either 'urgent' or 'non-urgent', with the first being collected in the first few minutes after a pulse and the other later in the day, during quiet periods. More solutions (now with ITER also in mind) like data compression and selective acquisition rates are also being studied and may be implemented in a near future.

1.2 Wendelstein 7-X

The Wendelstein 7-X (W7-X) reactor is an experimental stellarator built in Greifswald, Germany, by the Max Planck Institute of Plasma Physics (IPP), and completed in October 2015. It is currently the largest stellarator in the world and it was designed to operate with up to 30 minutes of continuous plasma. Its first plasma was created in December 2015 and regular operation has yet to start.

The CODAC of W7-X is based around a modular system called Control and Data Acquisition Station (CoDaStation). It was designed with the specifics o W7-X in mind and prototypes were tested on other installations (like WEGA stellerator) before deployment. The core of the CoDaStation software is a generic framework that solely provides the necessary infrastructure to acquire and process signals. In order to achieve high modularity, there are many components on the system, each with well defined functions nad interations. On the data acquisition side there is:

- Buffers: signal packages containing data with the same *time group* (similar time-stamp)
- Signal providers and consumers: generate a signal and recieve one or mores signals, respectively
- Routers: transfer data from *signal providers* and *signal consumers*
- Routing jobs: define if routers act synchronously or asynchronously, according to urgency or special requirements

On the control side there is:

- Properties: the state of a CoDaStation component
- Controllables: retrieve *Properties* from the CoDaStation and translate them into component-specific actions
- Station states: define the overall state of the CoDaStation
- Controller: interface for external control systems to issue control commands

These come together on packages called AddOns, which are used to integrate subsystems thus providing CoDaStation with new functionalities. These can have *controllables*, *signal providers* and *signal consumers*, interacting with the main system through a fixed software interface. Thanks to this, new features can be easily tested outside W7-X and also CoDaStation components can be replaced by mockups. There are also *Supervisors*: software that operators use to monitor *Station* and *AddOn* states and act if necessary. This can also be done remotely, thakns to JXI based supervisors.

1.3 NIF

The National Ignition Facility (NIF) is a large laser-based inertial confinement fusion research device, located at the Lawrence Livermore National Laboratory in Livermore, California, in the United States of America. Construction began in 1997, being ready for operation in 2009 and by 2010 the National Ignition Campaign started with the goal to produce more fusion energy than deposited on the target. However the campaign ended in 2012 without reaching its main objective and currently NIF is mostly devoted to materials studies.

At the heart of the control and data acquisition system of NIF is the NIF Data Repository (NDR). The NDR integrates many autonomous database systems thus being called a federated database. It allows the collection of all information relevant to the experiment, specifically campaign plans, machine configuration, calibration data, raw experimental results and processed results. There are three main events that trigger data collection and automated data analysis: laser alignment, shot time capture, and post shot optics damage inspection. Recent data is stored on-line, least recently used data is migrated to near-line storage and older data storage is subjected to different policies depending on the type of data. In 2009 NIF was generating approximately 66 TB of data per year, most of them 2-D images.

NDR uses industrial and scientific standards for its operation, such as:

- Web Services Business Process Execution Language (WS-BPEL) for preparing shots and organizing data collection, transfer and analysis;
- HDF5 files for storing calibration, background and diagnostic data;
- Uniform Resource Names (URN) to identify data and accessing it through Web services.

One of the goals of NDR is foster collaboration between universities and research institutes. To achieve this, all data and other content can be shared on the Web, with researchers being able to decide what they want to share with everyone, a restrited group of colleagues or to keep private.