THE PRESENT STATUS OF THE NAC CONTROL SYSTEM

H.F. WEEHUIZEN, R.K. FISCH, I.H. KOHLER, M.E. HOGAN, P.J. THERON

National Accelerator Centre, P O Box 72, Faure, 7131, Republic of South Africa

A microcomputer based distributed control system has been developed for a 200 MeV cyclotron and two injector cyclotrons. The system is based on some 25 PC-compatible microcomputers grouped into functions of operator consoles, instrumentation controllers and database nodes communicating over an Ethernet LAN. The system has been operating and steadily expanded over the last two years to replace the original system which is close to the end of its useful life. This paper looks briefly at the design of the system, discusses the success of the integration process, difficulties encountered and the advantages it offers over the old system.

1 Introduction

The National Accelerator Centre operates a 200 MeV Separated Sector Cyclotron and two injector cyclotrons[1]. One of the latter is a light ion injector while the other has two external ion sources. The one is an ECR source which is used to accelerate heavy ions while the other is a polarised ion source. The original control system used to control this facility is of classical design and consists of a minicomputer controlling remote hardware which has little or no intelligence. This system performed adequately for the first seven years of operation of the cyclotrons before all the physics research facilities were complete and the second injector cyclotron was operating. However this system has aged and the requirements of the facility began to outgrow the capabilities of the minicomputer system.

There was thus a need to develop a new system that can cope with the anticipated future requirements of the NAC facility. These are briefly as follows:

(a) It must be readily expandable to at least double the capacity of the original system.
(b) It must be sufficiently flexible to cope with the changing requirements of a partly experimental facility such as at the NAC.
(c) Hardware failures should have a minimal effect on the system.

A distributed system offers these advantages as new or extra functions are usually handled by extra nodes. It is also possible to easily reconfigure the system to cope with changing system requirements. Loss of a node means only the loss of the function that that node provides. This system is described in the next section.

2 The Distributed Control System

In keeping with control strategy, this system is designed to be as homogeneous as possible. It has five separate functional parts namely, network, instrumentation nodes, console nodes, database nodes and graphics nodes. Its logical layout is illustrated in figure 1. The design choices for this system are discussed in more detail in reference 2.

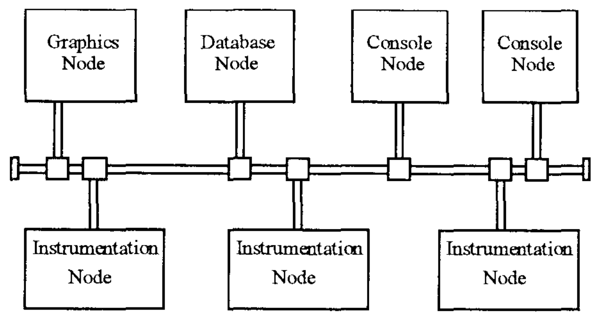

Figure 1: A block diagram of the distributed system.

2.1 The network

This is the Ethernet CSMA/CD 10BASE5 network which was chosen because the product was mature and readily available. It is frequently rejected for use in the control environment because it has a probabilistic transmission delay rather than the deterministic delay of the token passing protocols. In this application the loading is kept light to avoid this becoming a problem and the network performs very well. Currently the average loading is about 2,5% with the peak loading at about 18%. The average rate of collisions is about one every 10 seconds, so the probabilistic component is negligible. At this loading the transmission delay of the Ethernet system is considerably less than that of the token passing systems, thus the chosen system offers a superior performance[3].

The communication strategy is designed to keep the traffic on the network low. In event driven systems, communication is initiated every time a state change occurs, thus traffic is proportional to the number of events occurring in the system. Unfortunately when an emergency situation develops, a large number of state changes can occur therefore generating a large amount of network traffic. Transmission delays could then become significant at times when good network response is needed. Thus a combination of event and state driven philosophies is used. All operator initiated traffic is event driven giving a quick response from the system. The loading from this is not heavy as the rate at which a human operator can generate commands is very limited compared with that of a computer system. The actual value readback is partly state driven, partly event driven. If events at a node occur at intervals between 1 and 15 seconds, then the communication is event driven. At intervals less than 1 second, the rate is limited to once a second, thus becoming state driven. At intervals greater than 15 seconds, a node is forced to transmit the status of all its variables, thus indicating to all the other nodes that it is still active.

2.2 The system nodes

The decision to use PC compatible computers for the nodes has proved to be good . The costs have remained low despite a dramatic increase in the performance of the hardware and they have proved to be reliable. A mix of 386 and 486 based computers are used. It has been found necessary to have 16 Mbyte of memory in the computers as the networking software utilises a fair amount of memory and performance becomes unacceptably slow with less memory. OS/2 continues to be used as it has matured and is now a stable and reliable operating system. LAN Manager is still used with the database server although this will be changed in the future as the latest version of SQL Server is no longer supported on OS/2.

2.3 The instrumentation nodes

These interface to the electronics that control or measure the cyclotron parameters. Two interface types are used. The one is the CAMAC system while the other is an expansion I/O bus that was originally developed as a national 8 bit processor bus. It is used as a communication medium between the PC computers and intelligent I/O interfaces based on the 8051 microcontroller. The latter is the slightly newer system and was introduced for economic reasons.

The function of the instrumentation controller is centred around a variable table which holds all the values of the parameters controlled by the node. There are several values for each parameter. The first group of values consists of a set point reference value and an associated reference status. The second group holds the actual value and its status, including an error status if relevant. Further fields are used for housekeeping and linkage requests to the variable.

What has been described so far forms the standard control system software for each instrumentation node. It communicates with the network in a standard manner and provides a standardised interface to all the application software. The latter take the cyclotron related parameters and translate these into the device specific operations to effect the control of the hardware. Application tasks are activated when any change to the reference group occurs. Those tasks providing the actual value information are scheduled on a regular basis to update the actual value data in the table.

A communication task transfers information between the network and the variable table. This spontaneously transmits the table contents over the network. As described previously, the repetition rate of this transmission is limited to be between once every 15 seconds and once a second.

The communications task also receives commands from the console nodes to control the reference value of the variables. Before a reference value can be changed, a request must be made to be allowed access to the variable. Requests can come from several sources but only one may control it at any given time, so any contention for access must be resolved. Because the request/grant action must be atomic, the only effective place for this to be done is in the process that supervises the table. Thus the source that first makes the request is granted access and all subsequent requests are refused until the initial source relinquishes control.

Functions such as the acquisition of dynamic data would be handled by a separate process as these would be application specific.

2.4 The console nodes

The video screens of the original control system display 20 cyclotron variables at a time, each set of variables displayed being called a page. Each variable displayed consists of the variable description, its reference value, its actual value and status. Operator controls consist of a reduced function keyboard, two reference input joysticks called set point units and a programmable touch panel interface. Dedicated keys on the keyboard allow the latter three devices to be linked to any displayed variable. The keyboard is also used to select the page required for display. The set point units provide infinitely variable control of the selected variables while the touchpanel is used to select predefined states. While this style of console is effective, the original implementation was particularly cumbersome, thus obviating the use of the portable consoles.

The design of the new consoles had to satisfy three criteria.

(a) They have to use the standard PC interfaces thus needing the minimum amount of specialist software to support them, providing the fewest restrictions on the use of standard commercial software.
(b) They have to provide facilities as least as good as those of the original consoles.
(c) They have to be transportable so that a console can be placed at any reasonable location for development or maintenance purposes.

Using a mouse as an input device and the flexibility of the graphical user interface of the OS/2 operating system, it was easy to reproduce the functionality of the original console. As discussed later, it was easy to provide a far superior interface for the display of dynamic data on the new console. However, controlling the standard variables using a purely mouse based operator input is inconvenient. Page selection, which necessarily has a tree structure, is tedious and slow. Also mouse movement relies on visual feedback for correct operation. The main console has several other instruments which are frequently watched by the operator while parameters are adjusted on the console, therefore the visual feedback is broken while a variable is being adjusted, resulting in erratic and erroneous adjustments. Thus a reduced function keyboard and joysticks are used at the main console. The former is used for rapid page selection while the joysticks allow predictable control of cyclotron variables.

Two further features have been added to the new console. On the original system, the value of a variable could only be adjusted by longitudinal movement of the joystick. Forward movement of the joystick increases the value of the variable while backward movement decreases it. On the new system this is also possible but lateral movement of the joystick changes the decimal position of the cursor within the variable. Thus leftward movement of the joystick increases the increment value by a factor of ten for every digit moved and *vice versa* for rightward movement. Alternatively the other new feature can be used which is keyboard data entry. Page selection always involves an alpha numeric entry with the first character being alphabetic. The new interface is designed such that if a numeric entry is made on the keyboard, it is interpreted as a new value for the currently selected variable.

Part of this new operator interface is a LCD touchpanel interface which emulates the operation of the original CRT based touchpanel. The former uses infrared beams as the sensors for the programmable keys.

All the keyboard and joystick interaction with the PC takes place via the standard keyboard interface while the communication with the touchpanel is split between the keyboard interface and the parallel port. The graphics output to the LCD display takes place via an in-house developed driver controlling the parallel port while operator input is interpreted as standard keyboard codes and forms part of the reduced keyboard interface.

This new operator interface is only used at the main console where good flexibility and efficiency are required. Portability is not a requirement. The portable consoles are entirely mouse based. As there are no other instruments to be observed, the mouse based interaction is satisfactory, although keyboard control is still possible with the standard keyboard.

The console software is designed around a large variable table which holds a copy of the information in the instrumentation nodes together with the addresses of the nodes from which they originate. The information in this table is updated by the transmission of the variable data from the instrumentation nodes. When a page is selected for display on the console information is selected from the local table process from where the commands are directed via virtual circuits to the correct instrumentation nodes.

At startup the page layouts and variable groupings are fetched from the database node and held in memory thus forming a local memory resident database. The latter is used in preference to a networked or disc based database in the interests of speed.

2.5 The database node

The NAC has had a long period during which there has been construction, development and modification of the facilities on the site. Therefore it is an important criterion that the control system be flexible and readily expandable. This is most easily met by using a database to hold the configuration information for the control system. Thus it is possible to reconfigure the parameters of the system at a single point to allow for expansion and modification.

When the system is started, subsets of the main database are transferred to each node of the system as required and remain memory resident for access by the local node. The console nodes each have a complete set of the page configurations and display formats while the instrumentation nodes have a list of variables pertinent to their respective nodes.

Another function of the database node which is still to be developed, is to log the status of the system at regular intervals for collecting trending information and for system status recovery after interruptions in operation or return to operation at a previously used energy level. The use of the database node is such that it is not crucial to the operation of the system. If it did fail, operation of the system would continue but several important facilities would be lost. Thus a second node is available on standby if the primary node should fail.

Currently SQL Server is used as the database server running with LAN Manager to provide the network interface. As this is no longer supported under OS/2 it is planned to replace this with DB2/2. This is fully network aware and so requires no supporting network software.

2.6 The graphics node

This is a specialist operator interface that was designed primarily to provide a dynamic display of the beam profiles as derived from profile grids (harps) and profile scanners. The latter devices are situated on the various beamlines throughout the facility. Local instrumentation nodes control these devices and acquire the raw data from them. Each scan derived from the profile devices is sent as a packet to the graphics node where it is processed and presented for display. Virtual circuit communication is used to provide flow control over the network as the graphics display task is computation intensive and is thus slower than the acquisition task.

This node provides a graphical representation of all the Faraday cups, neutron shutters and the profile monitors in the beam lines. Their positional status is displayed on the monitor, any error conditions being indicated by the colour of the background. Further status information can be obtained by means of a pop up menu. The position of any device can be controlled by the mouse. When the profile monitors are inserted the data acquisition system is started and the profile is displayed in the top two thirds of the screen. Up to six devices can be displayed simultaneously. Hard copy of the dynamic displays can be made on a laser printer.

This system has been very well received by the operators because it gives a clear indication of the status of the devices at a glance and the control of the devices is intuitive. The performance of the system both in terms of response to operator commands, indicated response and refresh rate of the dynamic display is far superior to that of the original system. The refresh rate when two or less devices are displayed is such that a delay has been inserted between scans to slow down the display to make it more readable for the operators.

3 Integration

Installation of the new system had to be done during scheduled service periods. The only suitable times available are a three week shutdown period in January and a one week shutdown period in July. Thus only a section of the system could be integrated during each period. The process was divided into six stages which took three years to complete. Control has remained available at all times on the old system while the new system was integrated. Where CAMAC is used, it was easy to arbitrate between the old and the new systems at the CAMAC level as the standard supports more than one source of computer control. The newer interface does not have this facility, but it requires a local processor to interface to the old system via CAMAC. Thus the local processor was replaced by the instrumentation node of the new system. The node was then also interfaced to CAMAC, retaining the link to the old system.

Except for the instance of dynamic data acquisition, no attempt was made to avoid control conflicts between the old and the new systems. Few problems arose as a result of this. Those that did were secondary effects and were eliminated by disabling control of those functions on the old system. The only instance where the complexity of control was such that the facility had to disabled on the old system before commissioning on the new, was that for the control and display of the beam profile information. The process of changing over was very simple and took only five minutes to effect.

The operators have experienced little inconvenience in the installation of the new system. However considerable planning had to be done by the control division and many hours of overtime had to be spent commissioning the new hardware and software in the limited time available during each shutdown. Presently the original system still functions but is due to be used only as a standby system in a month or two. It will be decommissioned once confidence has been gained with the new system.

4 Conclusions

The new distributed control system has proved to be successful. The use of personal computers throughout the system has been justified as they have proved to be reliable and extremely cost effective. The chosen operating system has also proved to be reliable, most problems arising from lack of familiarity and lack of adequate time for fully testing the application software owing to very severe time constraints during installation. The system has proved to be flexible and easily expandable with no deterioration in system performance as the system has expanded.

Machine parameters are readily available for system analysis and diagnostic purposes. Specialist programs for machine optimisation can now be readily interfaced, saving significantly on development time.

References

1. G F Burdzik *et al.*, *Proc. 11th Int. Conf. on Cyclotrons and Their Applications* (Ionics, Tokyo, 1986) pp. 422-425.
2. H F Weehuizen *et al.*, *Proc. 13th Int. Conf. on Cyclotrons and Their Applications* (World Scientific, Singapore, 1993) pp. 656-659.
3. H F Weehuizen *et al.*, *Proc. 7th IFAC Workshop on Distributed Computer Control Systems* (Pergamon, London, 1987) pp. 73-77.