# Future data acquisition at ISIS

*W. C. A. Pulford, S. P. H. Quinton, M. W. Johnson and J. Norris*
Rutherford Appleton Laboratory
Chilton, Didcot, Oxon
UNITED KINGDOM

## Introduction

To some extent at ISIS we are becoming the victims of our own success. Over the past year ISIS beam intensity has increased steadily to 100 microamps during periods of good running. With the instrument users finding it comparatively easy to set up data-collection runs, we are facing an ever increasing volume of incoming data. Table 1 illustrates this point by showing the data volumes collected by the scheduled instruments during the weekend of September 10 and 11, 1988.

Note that the figures for HRPD do not include contributions from the now functional 90° detector bank, which might easily triple the data volume from this instrument. Moreover, the problems involved with digesting this data input will be exacerbated by the introduction of new instruments, such as the single crystal diffractometer (SXD) and the high intensity powder diffractometer (HIPD, still at the design stage)– both of which have large area detectors, up to 100,000 elements in the latter case. The data format of these instruments has been discussed in a paper delivered by M. W. Johnson at ICANS-IX. Greatly improved detector technology, mainly involving large areas of zinc sulfide phosphor, are expected to contribute much to the capacity of these instruments as well as provide an enhancement path for many of the existing ones. It is clear that we are fast reaching the point where if we continue to use our current technology data collection techniques, our computer systems will no longer be able to migrate the data to long-term storage, let alone enable their analysis at a speed compatible with continuous use of the ISIS instruments.

## Overcoming the data volume barrier

Before describing the methods that we will employ to overcome the data problem, I will briefly refer to our current data acquisition electronics (DAE 1) and migration path. These are illustrated schematically in Figures 1 and 2. This system does work extremely effectively, but experience has indicated a number of inherent difficulties:

(a) Seven instruments are still equipped with VAX 11/730 computers as their Front End Minicomputers (FEM). Unfortunately, these machines usually possess insufficient processor power to perform some of the more complex initial data reduction. This frequently means that the raw data have necessarily to be networked to the HUB computer before analysis may be performed. Currently, we send all data to the HUB, anyway, as a matter of policy.

**Table 1** Data volumes at ISIS for 10-9-88 and 11-9-88.

| Instrument | Type of instrument E-elastic I-inelastic | Number of datasets | Approx. size of datasets | Total data volume(blocks) |
|---|---|---|---|---|
| HRPD | E-High resolution powder diffractometer | 119 | 3,000 | 390,000 |
| HET | I-High energy transfer spectrometer | 17 | 3,100 | 54,000 |
| LOQ | E-Low angle spectrometer | 15 | 4,400 | 65,700 |
| IRIS | I-High resolution inelastic spectrometer | 5 | 4,400 | 22,000 |
| TFXA | I-Time focussed analyzer | 3 | 650 | 1,950 |
| CRISP | E-Neutron surface reflectometer | 25 | 50 | 1,350 |
| LAD | I-Liquid and amorphous diffractometer | 1 | 1,570 | 1,570 |
| POLARIS | I/E-Polarized neutron diffractometer | 4 | 1,000 | 4,000 |
| | Total | 540,570 (270 Mbytes) | | |

b) The size of the bulk store memory, in which histograms are stored, is restricted to 16 Mbytes by the 24-bit address field of the Multibus in the system crate. This limits the complexity of experiments available to the instrument scientists.
c) The DAE error detection and analysis system of the FEM is crude and renders the process of monitoring and correcting hardware errors non-straightforward.

It is clear that the most effective method to improve on this situation is to reduce the data volume flowing between the DAE and the FEM and to provide facilities to monitor data acquisition within the DAE. For these purposes we must incorporate processing power closer to the point of data collection. Our preliminary thoughts as to how we might proceed are set out in the ICANS paper previously mentioned.

## The next generation data acquisition electronics (DAE 2)

We have now decided to implement the processing elements within DAE 2 in the form of intelligent memory boards illustrated schematically in Fig. 3. The CPU

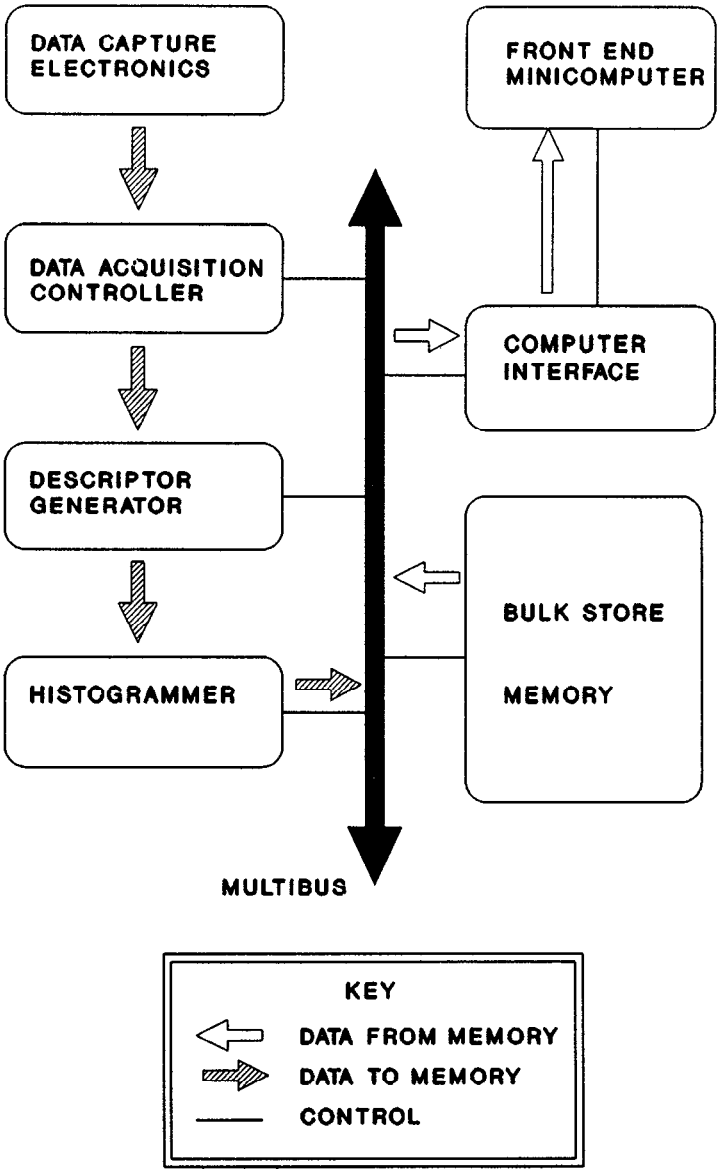


**Fig. 1** Existing data acquisition electronics.

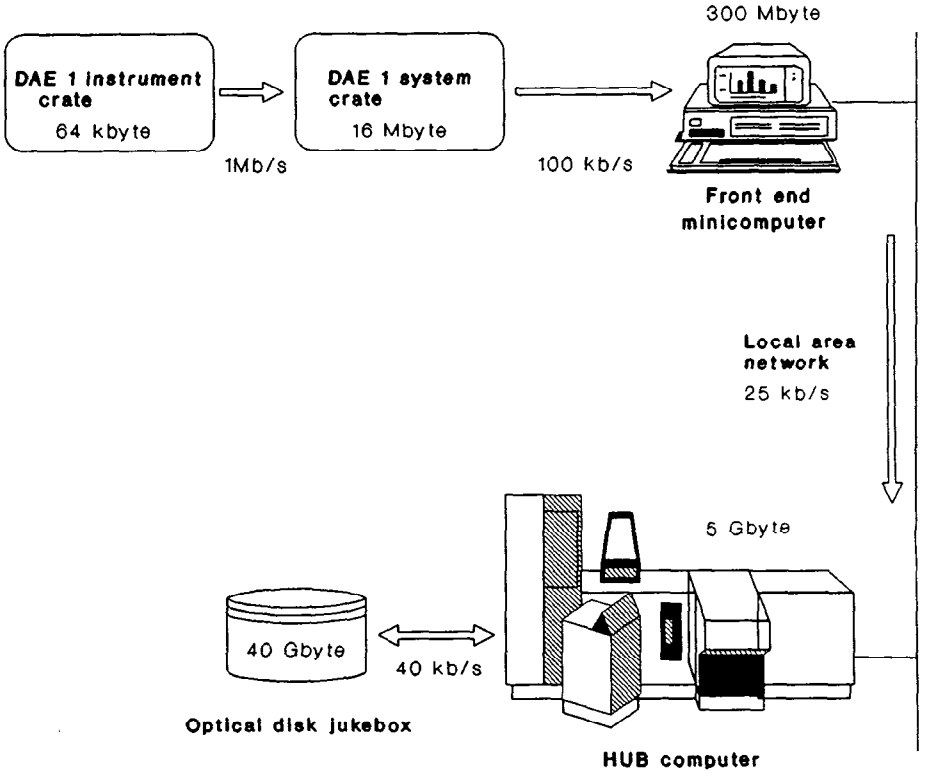


**Fig. 2** Current ISIS data migration path.

engines to be employed are TRANSPUTERS rather than Motorola 680x0 or National Semiconductor NS32032, mainly because we have not been able to find a commercially supported multiprocessing and multiprocessor operating system for the latter serial processors; whereas, from the second quarter of 1988, a variety have been available for the T4 and T8 transputers. Moreover, transputer architecture automatically lends itself to increasing overall processing performance by simply configuring further transputers in parallel using the links. Despite the flexibility of the transputer-link architecture, it was deemed necessary that the DAE 1a remains based on a databus because the 300 ns read/modify/write cycle of the T800 and the peak link speed of 1.5 Mbytes/second would impose too great a practical limitation on the allowed data collection speed should only the links be available.

The major enhancements desired for DAE 2 compared with DAE 1 are as follows:

a) The address bus should be at least 32 bits wide to enable enough memory to be configured for all experiments in the future.
b) The databus components should be capable of read/modify/write cycles at greater than 4 Mhz in anticipation of future data rates in excess of the current value of 1 Mhz. The higher bus speed would also prevent online

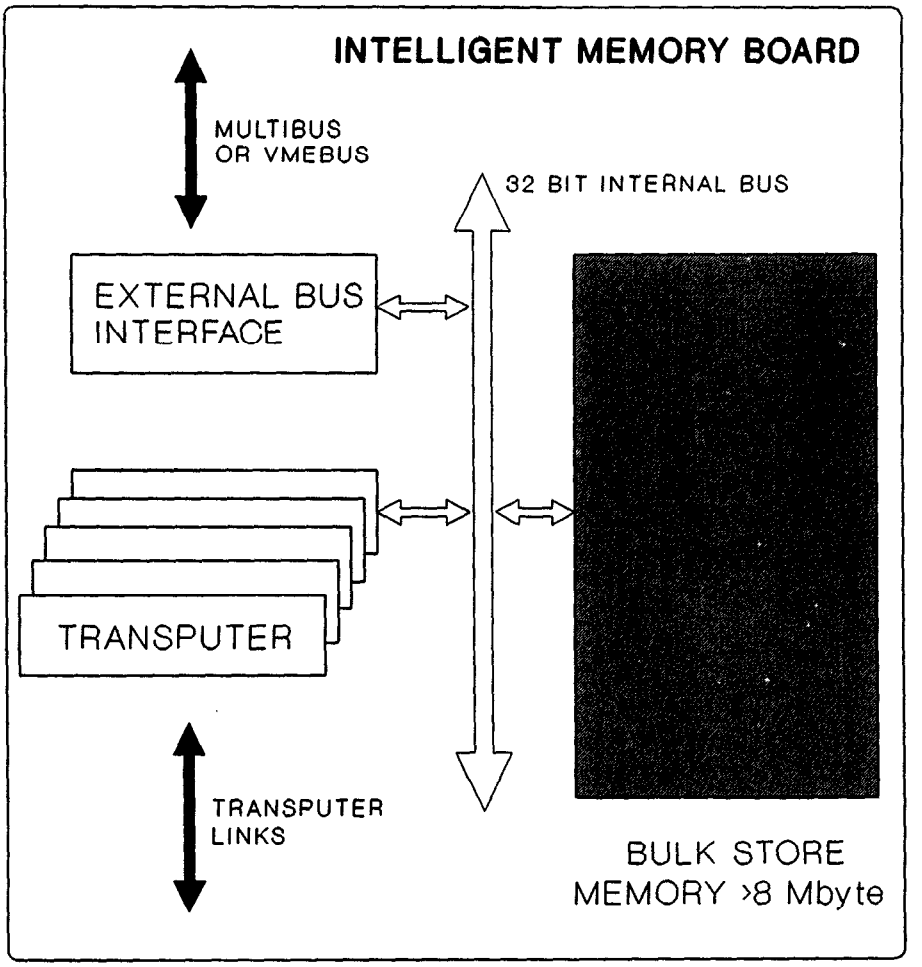


**Fig. 3** New intelligent DAE board.

data analysis from competing with data collection for the available bus bandwidth.

c) It may eventually be necessary to install a software programmable unit (descriptor processor), which maps the time of arrival and the position on the detector of a neutron event to the memory location to be incremented. This is occasioned by such instruments as HIPD, which have the capacity to map uneconomically large amounts of memory using the current relatively inflexible descriptor generator technology. Plans for the descriptor processor unit are not far advanced at this point, but it is reasonable to expect to use digital signal processors such as the Motorola DSP 56000, which are efficient at performing the type of multiply/accumulate functions we will require.

d) Scientists increasingly need to inspect the data as they are accumulated, particularly with reference to finding peaks and assessing the optimum duration of a run. With this in view, it is planned to incorporate a real time graphics display within DAE 2.

These requirements lead to the probable future DAE layout illustrated schematically in Fig. 4.

## Our current evaluation systems

In implementing DAE 2, we are attempting to preserve as much of our original intellectual and financial investment in DAE 1 as possible. Consequently, for the moment we have concentrated our efforts primarily on the system crate. This crate is the part of the data acquisition system interfaced to the FEM that controls data collection processes, 40-bit neutron event descriptors from the instrument crate, and histograms them in bulk memory. Our first evaluation system, known as DAE 1a is illustrated schematically in Fig. 5. Here we have made the minimum change necessary consistent with the overall plan since we have replaced the Multibus bulk store memory with in-house designed intelligent memory boards, each containing 1 Mbyte of RAM and one transputer. This enhancement is currently operational and will be available to ISIS instruments from the start to 1989. I expect this to be a major platform on which the necessary advanced software will be developed. We selected the CAPLIN QT0 as the interface between the transputer and the VAX mainly because it was the only one available that fulfilled our requirements. Nevertheless, this technically advanced board does have valuable properties, such as the ability to support three transputer systems, and a reasonably fast link to QBUS data rate. An interesting feature of DAE 1 is that it offers two routes for the transfer of data to the FEM, via the original computer interface (CI) or using the transputer link. The best attainable data rates for the two routes are 120 Kbytes/second for the CI and about 100 Kbytes/second for the link.

At the same time as developing DAE 1a, a closer approximation to DAE 2 is being assembled. It is shown schematically in Fig. 6. The VME bus was chosen for the following reasons:

a) It has 32 bit addressing and, in principle, supports bus speeds well in excess of the required 4 Mhz. The bus also supports read/modify/write as a single bus operation.
b) There is considerable commercial support for VME technology, particularly for the provision of functional boards. This may be important in the choice of graphics processors.
c) VME systems are much less expensive than the main competitor, which is FASTBUS.

The intelligent memory board used in this DAE 2 is the commercial INMOS BOO11 board, which has stations for three transputers but has only 2 Mbytes of onboard RAM. The shortage of RAM in the BOO11 or the need to reduce costs may eventually cause us to build an analogous board possible with fewer processors and at least 8 Mbytes of memory. Apart from this, the first major hardware design effort will be to produce a VME-based memory incrementer capable of receiving the output

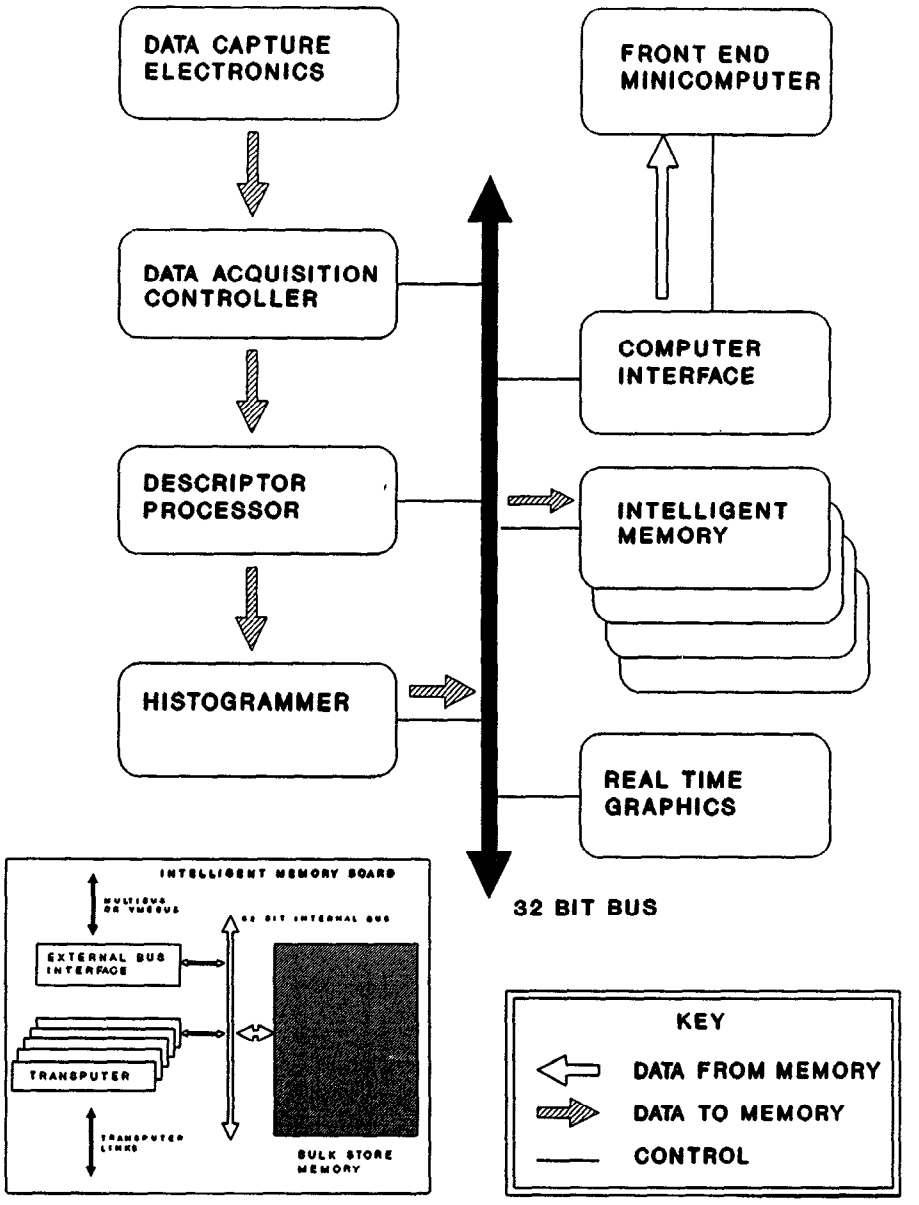


**Fig. 4** Probably future DAE layout.

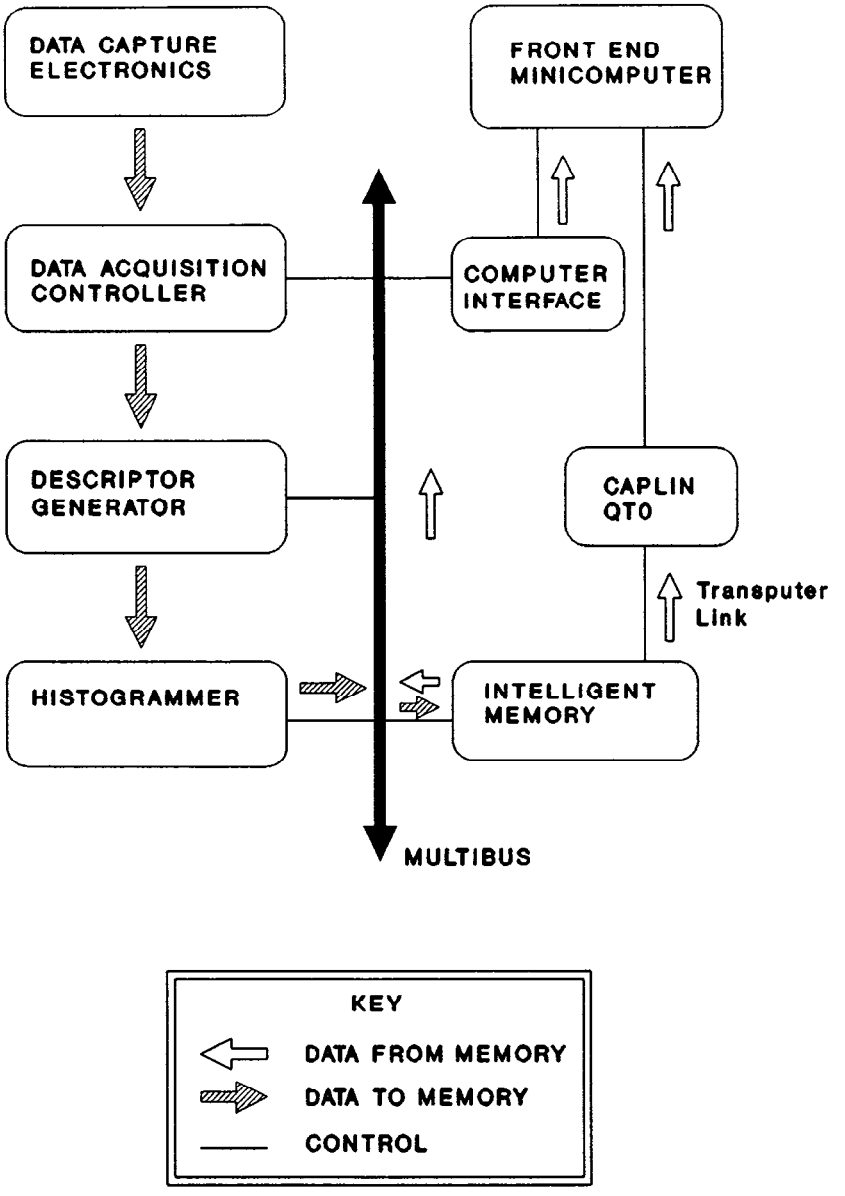


**Fig. 5** Evaluation system 1.

of the current descriptor generator. The future migration path away from DAE 1 to DAE 2 is clearly to reimplement the functional boards from the Multibus crate to the VME crate in bottom to top order in Fig. 6. This process may hopefully be accomplished gradually and with the minimum of disruption to the instruments' operation.

## DAE 2 software

We propose to write most of the software within the advanced DAE using a parallel implementation of the C language called multi-C. This supports multiple C processes programmed on an arbitrary configuration of transputers. The C language was chosen rather than the major alternative, OCCAM 2, for two reasons:

a) Many more of us are conversant with the language.
b) OCCAM 2 is generally supported by a system called TDS. We found this system to be extremely awkward to use, particularly from its normal host, the IBM PC. The folding editor, although elegant in concept, does not appear to us to be conducive in association with the language itself in producing major software systems.

In developing the software, we must pay great attention to trends within the computer industry, particularly with respect to user interfaces. In the systems area, it would be valuable if the DAE 2 were cluster sharable. An entertaining area might be to implement an MSCP analogue on the DAE 2 processor responsible for communications. (MSCP is Digital's proprietary mass-storage subsystem protocol.) To facilitate user applications for the DAE processing power, it may be valuable to implement at least part of the X-windows X protocol. This implementation would be supported by DEC themselves in the C language by the DEC-windows migration package.

The required software functionality changes, depending on the time domain of operation. While data are being collected, the main functions would be:

a) Supporting interactive peak searching for the purposes of monitoring data collection; checking instrument alignment and estimating optimum run duration.
b) Performing hardware-system-error checking. This would entail such activity as checking for time-dependent inconsistencies in the data as collection proceeds.

At the completion of a run, the emphasis of the software changes to compression and/or analysis of the collected data. In the cases of some instruments the data analysis doubles as a very powerful data compression. Two examples illustrate this point. On HRPD, spectra of adjacent detectors have similar profiles, but are phase-shifted by an amount dictated by the geometry of the machine. Large reductions in data storage requirements may be achieved by performing this phase shift using the DAE processors prior to data migration to the FEM. (This shift may eventually be accomplished using the programmable descriptor processor mentioned in Fig. 4.) With access to large amounts of processing power, it would also be possible to peak

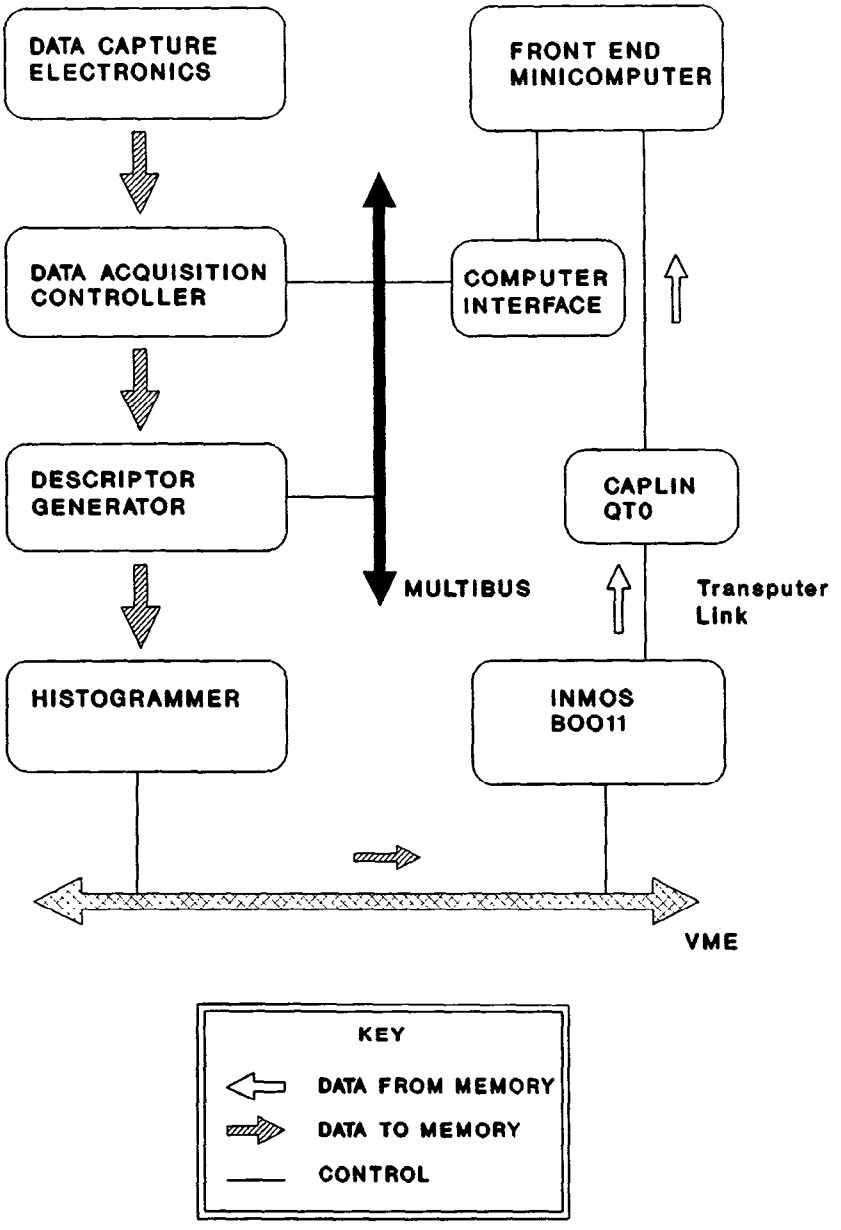


**Fig. 6** Evaluation system 2.

search the sparse data collected by a single crystal diffractometer and produce a potentially enormous reduction in data volume. However, even in less helpful situations, much compression may be performed by empirical methods, which make few prior assumptions about the data. The two most promising methods investigated so far are storing the data in a format analogous to a VMS descriptor and incremental byte packing. In the former case contiguous data points with the same value are not all stored but rather are abbreviated to a descriptor defining the data type (e. g., BYTE, WORD, LONG_WORD), a value count and a value field. The strategy may be extended to values which differ by less than the standard deviation from their mean, provided care is taken not to miss weak trends in the data. The second compression strategy depends on the difference between any two adjacent data points, not exceeding 256 for byte operations. Here only the initial value is stored in a whole word and the remaining data are stored as incremental bytes.

I expect the DAE 2 processing power to be used not only for the above-mentioned tasks, but also for analysis functions such as Rietvelt-profile refinements. To this end it is the intention to provide a set of software tools resident within the transputer domain but accessible to the instrument user, quite possibly via the popular GENIE data-reduction suite. This package is currently being rewritten to include better functionality, particularly with respect to access to windowing; hence the suggestion of an X-windows interface to DAE 2.