ANL-HEP-CP-96-82

THE ATLAS LEVEL 2 TRIGGER SUPERVISOR

Robert E. Blair, J. W. Dawson, and J. L. Schlereth
Argonne National Laboratory, Argonne, IL 60439 USA
Maris Abolins, Daniel Owen, and Bernard G. Pope
Michigan State University, East Lansing, MI 48824
Rodrigo Weber dos Santos
University Federal do Rio de Janiero, Rio de Janiero, Brazil

ABSTRACT

This paper presents an overview of the hardware and software proposed for the ATLAS level 2 Trigger ROI Builder/Supervisor. The essential requirements of this system are that it operate at the design Level 1 Trigger rate of 100kHz and that it support the technical requirements of the architectures suggested for the ATLAS Level 2 Trigger. Commercial equipment and software support are used to the maximum extent possible, with support from dedicated hardware. Timing requirements and latencies are discussed and simulation results are presented.

1. ATLAS TRIGGERING SCHEME AND THE ROLE OF THE SUPERVISOR

The ATLAS level 2 trigger will consist of a combination of general and special purpose computing systems connected to readout buffers (ROB) on the input end and a level 3 computing farm on the output end via high speed networks[1]. In order to keep the amount of data transferred into the level 2 computing systems low, only data associated with level 1 trigger objects, like jets or isolated electromagnetic clusters (electrons or photons), will be passed to the level 2 processors. This regions of interest (ROI) driven approach decreases the data which must flow at the maximum 100kHz level 1 accept rate, but places additional processing demands on the system to select for transfer to level 2 only the data associated with the ROI's. Information about the ROI's from the level 1 system will be assembled and organized by an ROI builder and then passed to the level 2 Supervisor which will coordinate the transfer of data from the ROB's to the level 2 system. A block diagram of how the ATLAS trigger might be implemented is shown in Figure 1.

The Supervisor is required to perform tasks in the following categories:

1) Level 1 driven processes
   a) receive the level 1 ROI data
   b) select the level 2 processors for this event
   c) send ROB dispatching requests

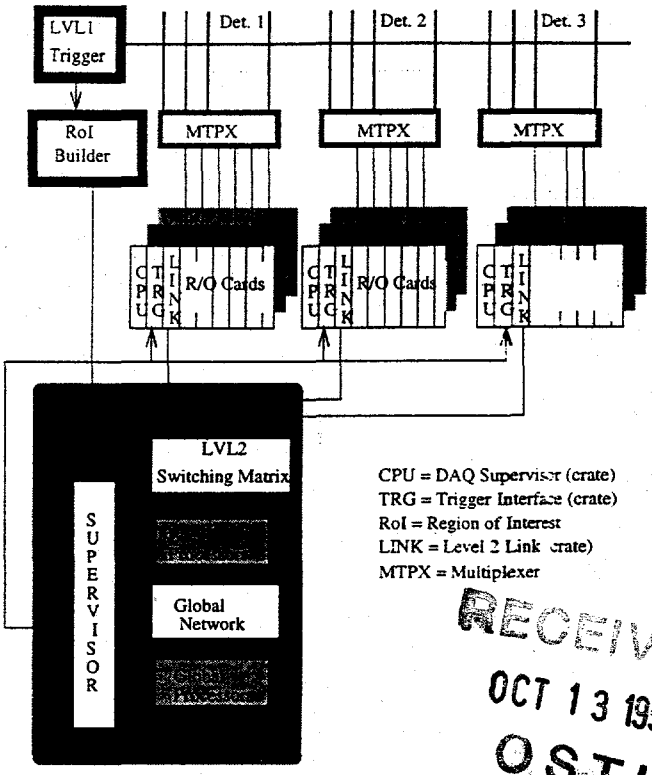

Fig. 1. ATLAS Level 1 and Level 2 Trigger

2) Level 2 driven processes
   a) receive the level 2 decision
   b) communicate the results to the ROB's to either drop the event or keep it for further processing

3) Non-data driven processes
   a) maintain list of allocated level 2 resources
   b) monitor system performance and report error conditions

2. SUPERVISOR DESIGN

The system envisioned to perform this set of tasks is shown in Figure 2. It is based on a distributed set of processors, housed in VME, with buffered I/O performed by dedicated hardware connected to each

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

processor via a dedicated PMC interface. One of the processors, called the Steward, plays a special role in managing resources and performing Level 2 driven tasks. It uses the VME bus for lower bandwidth interprocessor communication, and dedicated hardware to route ROI information to one of several processors used to perform the level 1 driven tasks.

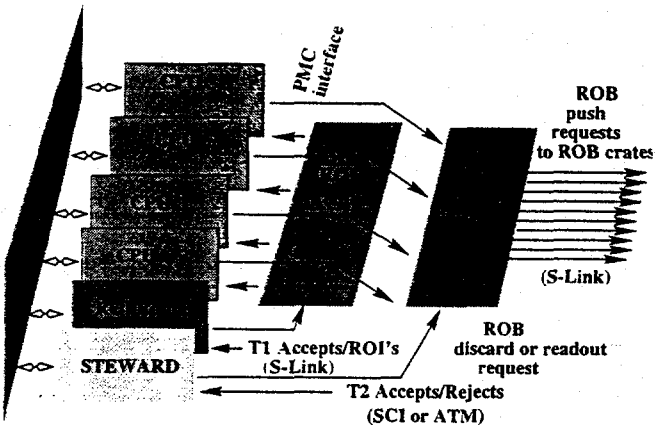

Fig. 2. Supervisor for ATLAS

To deal with the input rate, the supervisor uses an input router implemented in hardware. This hardware receives the incoming ROI information from the Level 1 Trigger and routes this information directly to one of the ROI Processors via a 32-bit wide FIFO on a PMC mezzanine card on the relevant ROI Processor. A round robin algorithm is implemented in hardware so that the router automatically, without intervention by the Steward Processor, gates the ROI information from the next Level 1 Accept to the next ROI Processor. Since the router's only function is to switch the incoming ROI information to the FIFO of the appropriate ROI Processor, it is inherently deadtimeless. The ROI information does, however, come more or less asynchronously from six Level 1 Trigger Processors. Accordingly, the input hardware must route by event number the six ROI fragments to the appropriate processor, while dealing with the fact that an ROI fragment received may relate to a preceding or following event.

The Steward Processor provides essential management functions for the supervisor. It receives the Level 2 accept/reject decisions through a PMC mezzanine card, keeps track of status of resources (for example, which global Processors are free and can be allocated to Level 1 Accepts), handles broadcasts to ROB's, and performs monitoring and error detection.

Every processor, including the Steward, has ports to the output hardware implemented on a PMC mezzanine card which houses a FIFO to receive data from the Processor. When the ROI Processor has formatted the set of data to be output, that data is written directly to the output FIFO in the PMC card through the PCI port. No time is spent on bus arbitration in this process as the data are written directly to the FIFO's on the PMC cards.

The output hardware embodies a bus which receives data from the Processors' output PMC cards, and allows the data to be written to output channels. The bus and the PMC mezzanine cards in the Processors must support a simple arbitration protocol, performed by hardware on the ROI Processor's output PMC mezzanine card and the output hardware without intervention by the ROI Processors or Steward Processor. Once the bus is granted to one of the ROI Processors' output PMC cards, the FIFO contents are written to the appropriate output channels under hardware control by hardware on the output PMC card. As soon as it has written the output data for the event to the output FIFO, the ROI Processor is available for the next Level 1 accept. The further transfer to the output channels is handled by hardware at the output FIFO buffer on the PMC card.

An output channel could be connected by a fiber to any other element of the Level 2 trigger. We propose to use S-Link to transfer data from the Supervisor to the other elements of the Second Level Trigger, and also to bring ROI data from the Level 1 Trigger Processors via S-Link. The ROI Processors provide routing information in the form of flags and embedded addresses by which the routing hardware is able to route the data blocks to the correct output channels. Each output channel consists of a FIFO and the required hardware to drive the fiber or other transmission medium.

3. HARDWARE REALIZATION

There exists within the Level 2 Trigger community an effort (The Demonstrator Project) to implement in hardware subsets of the actual trigger architecture envisioned for ATLAS, and in fact there are currently three different architectures being discussed. Our objective has been to provide supervisor support for these three architectures with prototype hardware, striving to make our designs general enough that we may support these three architectures. Clearly, in a project which will be constructed several years from now, there is a danger in freezing technology choices too early. However, our feeling has been that a prototype hardware realization effort at this time is useful in leading to the judicious choice of architecture for the Level 2 Trigger and in providing insight into the bottlenecks and other problems to be encountered in the final design.

The prototype hardware we are developing consists of the input and output routers, PMC cards, bus structures,

and miscellaneous support hardware. We have produced PMC cards which basically satisfy the requirements of the supervisor support for The Demonstrator Project, and in fact for ATLAS itself. These PMC cards utilize the AMCC 5933 chip to manage the PCI port. Data rates as high as 65 MB/sec have been routinely achieved operating the PMC cards with a Motorola 1604 Processor using a 100 MHz PowerPC running under OS-9000, and data transfer seems quite robust.

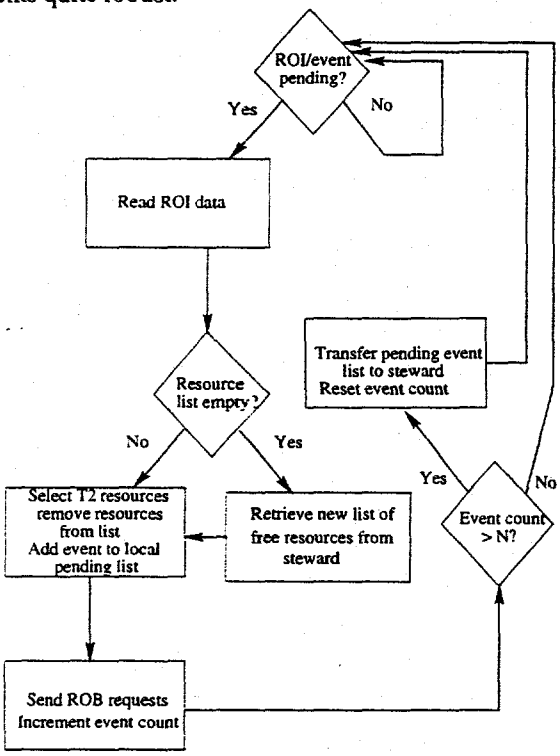

Fig. 3. Simplified ROI CPU Processing Loop

4. PERFORMANCE ESTIMATES

A very simplified flowchart of the polling loop for an individual ROI processor in the system described above is shown in Figure 3. A similar flowchart for the steward processor appears in Figure 4. Using transfer rates measured at CERN for the CES RIO2 processor for the PCI access and the VME transfers and estimates based on algorithm benchmarks done on other processors, we conclude that a small system of four or five processors should be able to manage the level 2 system at the required rate. These estimates, which do not include contention on the VME bus nor overheads associated with switching between transfer tasks and computing, lead to overall times for a single steward system of 8 microseconds per event for the steward and 8 microseconds for each ROI CPU to process an event.

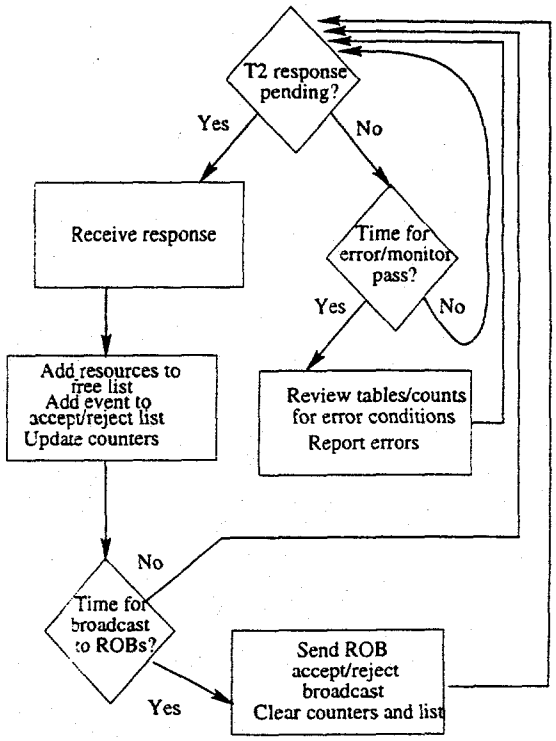

Fig. 4. Simplified Steward CPU Processing Loop

5. CONCLUSIONS

Our preliminary design for the Supervisor for the ATLAS Level 2 Trigger employs a farm of VME-based processors for processing ROI information, one or more processors for managing the Supervisor, fast dedicated hardware for receiving and distributing ROI information from the Level 1 Processors, and fast dedicated hardware for distributing the processed ROI records to the Readout Buffers or other elements of the Level 2 Trigger. There are presently three architectures being discussed in the Level 2 Trigger Community and, in some architectures, S-Link and/or SCI communication are replaced by ATM. All fast I/O with the processors is accomplished via PCI ports implemented with PMC cards.

Prototype hardware development to date has focused on the PMC cards, dedicated routers for input and output, and bus structures to support the data transfer. Initial results on PMC cards have indicated that we can achieve data transfer rates approaching 65 MB/sec with excellent data integrity.

REFERENCES

1. W. W. Armstrong *et al.*, "ATLAS Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN", CERN/LHCC/94-43 (December 1994).