ATLAS TDAQ RoI Builder and the Level 2 Supervisor system

R. E. Blair[1], J. Dawson[1], G. Drake[1], W. Haberichter[1], J. Schlereth[1],
M. Abolins[2], Y. Ermoline[2], B. G. Pope[2]

[1] Argonne National Laboratory, 9700 south Cass Avenue, Argonne, IL 60439 USA
[2] Michigan State University, East Lansing, MI 48824-1116 USA

Abstract

The ATLAS High Level Trigger (HLT) uses information from the hardware based Level 1 Trigger system to guide the retrieval of information from the readout system. The Level 1 Trigger elements (jet, electromagnetic, muon candidate, etc.) determine Regions of Interest (RoIs) that seed further trigger decisions. This paper describes the device - the RoI Builder (RoIB) - that collects these data from the Level 1 Trigger and the Level 2 Supervisors (L2SV) Farm that makes these data available to the HLT. The status of the system design and the results of the tests and integration into ATLAS TDAQ system are presented.

I. Region of Interest concept and RoIB/L2SV system prototyping

The Level 1 Trigger system (LVL1) identifies a number of RoIs for the HLT - spatially limited areas ('roads') in the detector with candidates for phenomena to be triggered. The information from Level 1 includes the selected trigger type and the details of where in η and φ the trigger objects (e/γ, μ, etc.) that caused the event to be accepted originated. Using this information as guidance, specialized algorithms request a sub-set of the event data from the ROSs to perform the event selection. In this way only a few per cent of the event data need to be transferred initially to the HLT system — thus considerably reducing the network bandwidth required.

The initial ideas and requirements are summarized in [1, 2]. The RoIB and L2SV system have been prototyped a number of times, have undergone a series of design reviews and tests, and have been integrated with individual parts of the Level 1 Trigger system and the HLT.

The Region of Interest Builder (RoIB), located in the underground counting room, takes raw event fragments from various Level 1 Trigger sources, assembles all the fragments of a given event into an RoI record, selects a target Level 2 Supervisor (L2SV) processor in a farm, and then sends the RoI record to this processor. From there the RoI records are distributed to the High Level Trigger (HLT) processors that require them for further event selection and disposition.

The Level 2 Supervisor farm, located in the surface TDAQ computing farm barracks, is comprised of commercial rack mounted PCs. They are connected to the RoIB via ATLAS standard S-LINKs and to the High Level Trigger via Ethernet. Each Level 2 Supervisor processor is thus subjected to a fractiono f the Level 1 accept rate avoiding issues of hight input and output bandwidth.

The overview of the RoIB/L2SV system is shown in Fig.1. The baseline implementation of the RoIB is a VMEbus system which includes a Single Board Computer (SBC) for configuration, control and monitoring. It is composed of two stages: input and assembly. The input stage consists of Input Cards that receive and buffer the RoI fragments. These cards subsequently send the RoI fragments to Builder Cards in the assembly stage where the RoI fragments are assembled into RoI records. Each Builder Card can service up to four supervisor processors. The number of builder cards within the system is not limited and is dictated by the rate that a supervisor processor can sustain.

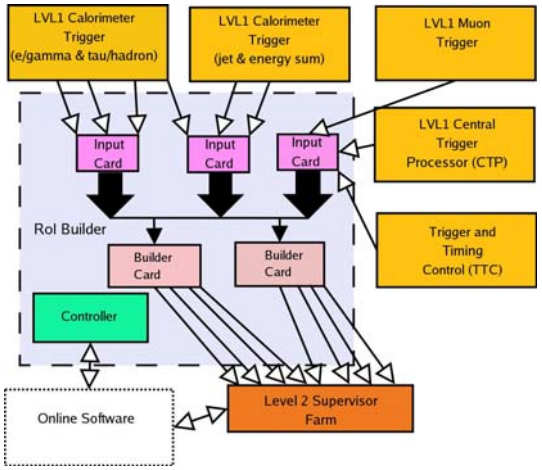

Figure 1: RoIB/L2SV system overview

A prototype of the RoIB was built and tested during the course of 1999. It was based on a pair of 12U VMEbus cards, an Input Card capable of handling six S-LINK inputs, and a pair of Builder Cards able to output to a pair of supervisor processes. This implementation utilized 76 Altera 10K40 FPGAs and 8 10K50 FPGAs. The system and early performance measurements are documented in [3].

Testing has shown that combining RoI fragments from several sources using an FPGAbased device is feasible and that a L2SV consisting of four 300 MHz Pentium II PCs is sufficient to receive the RoIB output rate of 100 kHz. Subsequent tests with prototypes of the muon Central Trigger Processor Interface and the Central Trigger Processor ROD modules of the LVL1 system have made a start on debugging the component interfaces and have further demonstrated that data input to RoIB could be achieved at the required rates [4].

II. RoIB/L2SV System Design

The final implementation of the RoIB cards will be on 9U VMEbus cards. Each of 4 Input Cards accommodates 3 S-LINK input LDC's (Link Destination Card), and can service up to 4 Builder Cards (limited by the custom backplane in the VME crate). Each Builder Card can service up to four supervisor processors sending data via S-LINK output LSC's (Link Source Card), The Single Board Computer (SBC) in the RoIB crate provides the configuration and monitoring information to the Input and Builder Cards.

A. RoIB Input Card

A simplified block diagram of the Input Card is shown in Fig. 2. All transfer of information from the Input Cards to the Builder cards is via J3 high density 250 pin connector and a custom backplane mounted in the rear of the crate.

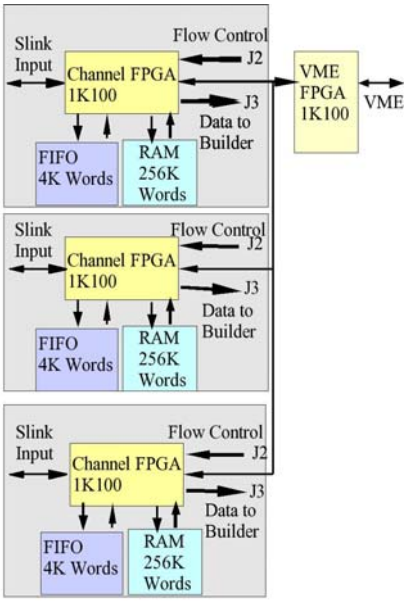

Figure 2: Simplified block diagram of the Input Card

Each Input card also has a diagnostic RAM initialized from VMEbus which allows emulating LVL1 fragments, and enables the operator to verify proper functioning of the RoIB system in a stand alone mode. The Input Cards have several modes of operation:

- Diagnostic Run Mode – the input data streams are provided by the diagnostic RAM's and the contents may be diagnostic data, test vectors from Monte Carlo or simulation results.
- Sniffer Mode – the RAM is written with Fragments from the incoming data stream from LVL1. The contents of the diagnostic RAM can then be accessed in block transfers from VMEbus.
- No Sniffer Data Mode – LVL1 data from an Input Card are transferred through the custom backplane to the Builder Cards with the 3 channels functioning independently of each other.

Each channel has an FPGA to provide the logic required for managing the data. Depending on the Mode that is defined by the status register, there are several paths that data may take. In Sniffer Mode and No Sniffer Data Mode the path of data is from S-LINK through the FPGA into the FIFO. In Sniffer Data Mode the S-LINK words are written to the Diagnostic RAM as they are read from FIFO. In No Sniffer Data Mode, if there are data in the FIFO and there is no Flow Control from the Builder Cards for this channel, data are read from the FIFO one word at a time, parsed and formatted into 20 bit words, and transferred to the Builder Cards. In Diagnostic Run Mode the path of data is from RAM, through the FPGA, into FIFO, from FIFO, through the FPGA, and after being parsed, through J3 to the Builder Cards.

The flow control signals are transferred via User Defined pins on J2. Since each Builder Card deals with input from 12 S-Link channels, it must provide 12 flow control signals. These 12 signal lines are bussed on J2 and are wire-or'ed. The first three of the flow control signals go to the first Input Card, the second three go to the second Input Card, etc.

There is an FPGA to handle the transactions with VME, which include 32 bit non-privileged transfers for reading/writing registers and 64 bit block transfers for reading/writing the diagnostic RAM(S). The VME FPGA also includes a number of registers, such as status, which are relevant to all three channels. A photo of the Input Card is shown below:

Figure 3: RoIB Input Card

B. RoIB Builder Card

Each data fragment contains RoI data collected from a portion of the Level 1 trigger system. The Level 1 information required for the HLT system is the concatenation of all such fragments for an event. This includes both the information about the trigger decision as well as eta and phi data for the subsystems that cause an event trigger. We refer to the collected RoI fragments for a given event as an RoI record, and to the subsystem on the Builder Card that builds the RoI record, as the Assembly Unit (AU, see Fig. 4).

The Input Cards pass RoI fragments to a set of Builder Cards. Each Builder Card communicates RoI records to four

Supervisor processors. The compiled RoI record is transferred to the target Supervisor processors using S-Link. Each of the Builder Cards is responsible for a subset of the events that trigger Level 1.

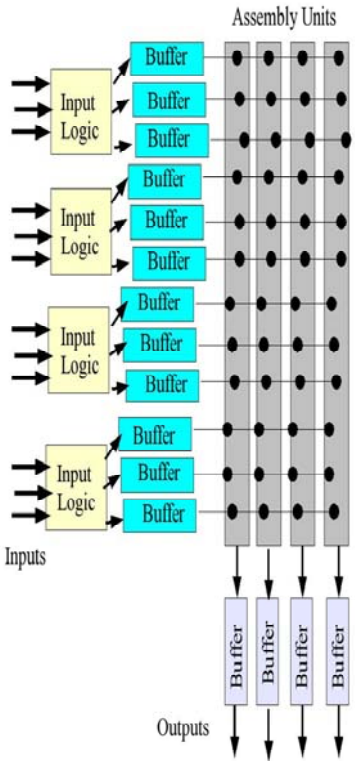

Figure 4: Simplified block diagram of the Builder Card

The RoIB uses basic round robin algorithm to assign Supervisor Processors. The system is expandable in units of four Supervisor processors by adding another Builder Card. The backplane is able to accommodate eight RoIB cards. Each Builder Card has registers which tell it which of the Level 1 channels are active, how many Builder Cards there are, which card it is in the sequence of cards, etc.

Events are allocated to Supervisor processors on a round robin basis, with the hardware automatically adjusting for the number of cards, number of Supervisor processors, etc. The firmware also allows the allocation of events to specific AU's based on criteria other that the Event ID, for example the Event Type.

Flow Control exists at a number of points on the RoIB, and is not treated as one continuous signal. For example, there is a Flow Control between the input FIFO in the Builder Card and the Input Card, Another example of Flow Control on the Builder Card is when the output FIFO goes half full it exerts Flow Control back to the builder logic.

The individual ROI fragments can be as long as 128 S-LINK words including headers and trailers, and are in the S-Link format. It is necessary to accommodate the time skew of arriving fragments, and accordingly a timer is started at the arrival of the first fragment of each event. If all the fragments have been received before the timeout, the compiled record is transferred to the target Supervisor process. If the timeout occurs the system transfers an incomplete record to the target Supervisor process. If a tardy fragment from a previous incompletely built record is received it is discarded.

The timeout and other parameters of the Builder Card are selectable from VMEbus as on the Input Card. The photo of the Builder Card is shown below:

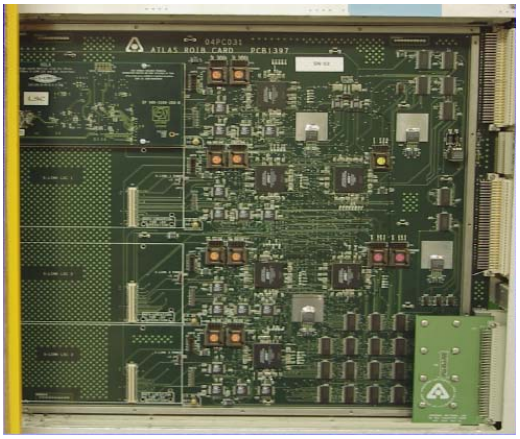

Figure 5: RoIB Builder Card

C. Custom Backplane

Signals are carried from the Input Cards to the Builder Cards via a custom backplane at 3.3 volt logic levels with special care taken to avoid crosstalk. The problem of driving data and clock lines was dealt with by reducing the clock frequency for the input cards to 20 MHz, by paralleling the drivers, and by limiting the number of Builder Cards to 4.

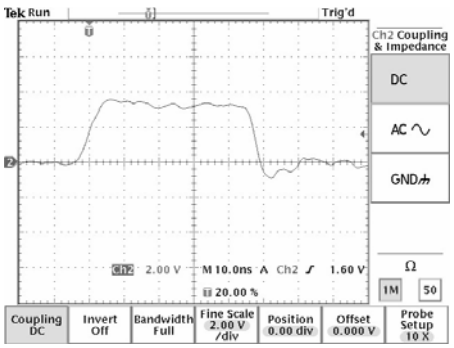

Figure 6: Data bit on the backplane showing rise and fall time

In order to check possible transmission errors, the firmware generates a checksum in the Input Cards on each fragment, transfer it to the Builder Card after each trailer, and compare it with a checksum being generated as the fragment is received. In the event of disagreement the Leve1 Event Number is written to a FIFO on the Builder Card. The photo of the custom backplane is shown on Fig. 7.

Figure 7: View of the fully populated RoIB crate

D. SBC and standalone tests

A VME processor VP110/010 from Concurrent Technologies is used as a Controller in the RoIB crate to access the VME registers available for programming and controlling the two types of cards.

It allows two levels of testing of the RoIB – basic single card initial checkout and multi-card integrated checkout (Fig. 8). Single card test for the Input Card includes single S-Link to sniffer RAM test and write/read RAM test, For the Builder Card there are spy FIFO and single S-Link output tests.

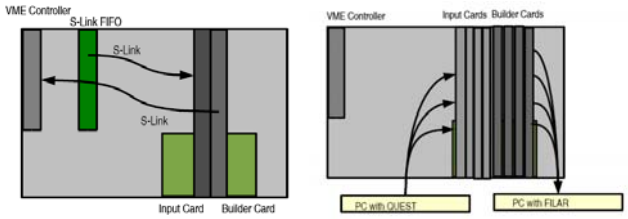

Figure 8: Basic and multi-card test setups

The RoIB test programs for the multi-card integrated checkout use external PCs for data sources/sinks and define 3 levels of complexity - spy FIFO (requires no external sources/sinks), check readout (requires external sink) and check readout with external data source and external sink, sample monitoring data.

The following conclusion about performance of the RoI Builder system was made based on the results of the standalone tests:

- throughput depends weakly on number of inputs due to parallel nature of system,
- throughput depends mainly on length of longest fragment,
- throughput of integrated system (external sources and sinks) was limited by readout rate,
- near 100 kHz is achieved for 3 inputs, 4 outputs.

E. L2SV Farm

The Level 2 Supervisor farm (up to 16 commercial rack mounted PCs) is connected to the RoIB via ATLAS standard SLINKs and to the High Level Trigger – via Ethernet. Each PC is equipped with FILAR card - quad S-LINK LDCs to PCI Interface. Each processor in the farm is responsible for distributing RoI records to the HLT processors farm and effectively managing the processing resources of the HLT farm via a load balancing algorithm. It receives the final decision on an event based on the result of the HLT processor. The decision results are communicated to the Data Flow Manager (DFM) so that accepted events can be further analyzed and rejected events can be flushed from the Readout System.

III. RoIB/L2SV Tests and Commissioning

The RoIB and L2SV system has undergone a number of prototypes, a series of design reviews and several tests, and integration with individual parts of the Level 1 Trigger system and the HLT.

A. H8 Test Beam

Initial test of a small RoIB system was in H8 test beam, it was integrated with Level 1 Trigger outputs from muon and calorimeter triggers. Control software was fully integrated with DAQ system..

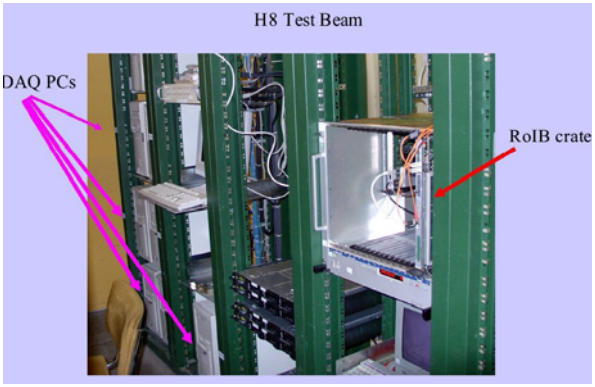

Figure 9: H8 test beam setup

B. RoIB in USA15

The commissioning stage has started with incremental installation of individual RoIB components and connection of the Level 1 Trigger system to a fully populated RoIB crate (see Fig.10).

The RoIB installed recently in the USA15 underground counting room contains 4 Input Cards and 3 Builder cards, however only 3 LVL1 Trigger inputs are connected at present and only 2 L2SV processors are used in the SDX1 counting room. The size of the L2SV farm will be augmented when dictated by increasing event rates during commissioning.

Figure 10: Final setup in USA15

C. TTC LDC

A mezzanine card for bringing TTC information into the RoIB system was developed. This card appears to an Input Card to be an S-Link LDC, but instead receives input from the TTC fibre into a TTCrx. The TTC information is latched on every L1 Accept, is formatted on the mezzanine card to resemble an RoI fragment, and is accepted and processed by the RoIB as if it were an RoI fragment. The data which are latched on a L1 Accept are the 24 bits of the Event Counter, 12 bits of the Bunch Counter, and 8 bits of Trigger Type. Once the data are latched, they are reformatted in the form of an RoI fragment, and written to a FIFO. If there is no flow control imposed, the pseudo fragment is next transferred to the Input Card through the S-Link port. An 8 bit register internally in the FPGA is incremented on Event Resets to provide the top 8 bits of the L1 Event Number.

IV. REFERENCES

[1] Y. Ermoline, The level-1/level-2 interface: RoI Unit, ATLAS DAQ note 94-34, 8 December 1994.

[2] Y. Ermoline, The level-2 trigger system supervision, ATLAS DAQ note 94-35, 8 December 1994.

[3] R. E. Blair et al., A Prototype ROI Builder for the Second Level Trigger of ATLAS Implemented in FPGA's, LEB'99, Snowmass, September 20-24 1999, ATL-DAQ-99-016, Dec 1999.

[4] B. M. Barnett et al., Dataflow integration of the calorimeter trigger CPROD with the RoI Builder and the ROS, EDMS Note ATL-DA-ER-0016 (2001).

[5] R. E. Blair et al., RoI Builder Manual, v3.6, ATL-DQ-ON-0005, 22 Feb 2005.