International Journal of Modern Physics A,

The D0 Level 2 Trigger

James T. Linnemann[a]
For the DØ Level 2 Trigger Group
Department of Physics and Astronomy, Michigan State University
E. Lansing, Michigan 48824, USA

The D0 Level 2 trigger (L2), a subsystem of the three-stage D0 trigger, reduces event rates from 10 KHz to 1 KHz. The system is organized as a global processor and preprocessors for the calorimeter, muon, tracker, and preshower subsystems. Preprocessors digest information from a single detector system and produce a list of the objects. The global processor receives the object lists, matches across detectors, and applies selection criteria to the objects. The system can buffer 16 events. The L2 design has been guided by extensive queuing simulations. The system is presently being commissioned, with all components available as either pre-production or production boards.

The basic parameter of the Level 2 (L2) trigger[1] for the DØ experiment at Fermilab is the input rate of 10 KHz, giving a nominal processing time budget of $100\mu s$. Within this time budget, a reduction of rate by a factor of 10 is to be accomplished, with good efficiency and introducing less than a 5% deadtime. The functions performed by the L2 trigger are collecting data from front-ends and L1 trigger processors, analyzing and formatting this data into components of physics objects, combining data across detectors into higher-quality physics objects, and selecting events on script-controlled criteria based on 128-bit L1 trigger decisions. The hardware aspects of the project are driven by the various input formats, and the need to apply processing power to the data. The software aspects are determined by the amount of processing needed for production of the physics objects, and the complex scripting mechanism needed to support up to 128 decision scenarios.

The architecture chosen involves a preprocessor for each subdetector system, and a global processor for integration of the data. The input cards of the system buffer up to 16 events (as do the DAQ front ends), so the preprocessor and global processor work as a loosely-coupled or stochastic 2-stage pipeline. Each stage of the pipeline has the same nominal time budget. Though queueing simulations have shown some mild time budget restrictions due to the two stages, fluctuations in the time to process individual events dominate the decrease from the nominal time budget ($100\mu s$) to the actual time budget ($50-60\mu s$) allowable before deadtime is incurred. The architecture was heavily constrained by a front-end readout requirement that decisions be reported in the same order as the events arrived.

[a]Support from the National Science Foundation and Department of Energy is gratefully acknowledged. It has been a pleasure to work with, and learn from, my many dedicated and talented colleagues who have contributed to this project. I mention especially Dan Edmunds, chief architect of the L2 trigger, who has taught me more about electronics than I've been able to remember.

This rendered unattractive a simple farm architecture, which might have assembled events and handed them out to independent processors.

Data comes into the L2 system in two basic physical formats. The first is 1.3Gbit/s G-link serial data on optical fibers, used by the tracking system, the calorimeter, and the L1 trigger framework. The second, a lower-bandwidth format used by the muon system, is 160Mbit/s serial Hotlinks on coaxial cable. Within the L2 system, we use for serial communication 160Mbit/s Hotlinks on the standard UTP cables used for ethernet wiring. For parallel communication we used a high-bandwidth 320MByte/s custom 128-bit bus (Magic Bus or Mbus) in VME crates built to the VIPA standard.

The first task is to get data into the right physical format. The coaxial Hotlink data is received on analog-only cable input converter (CIC) cards, amplified, standardized, copied, and sent to the destination processors as differential UTP Hotlink signals. Similarly, the G-link data is received, buffered (for 16 events), and speed-converted to differential UTP Hotlink signals by Fiber Input Converter (FIC) cards. The fiber receiving is performed on a transition module (VTM) located at the back of the crate; this same transition module is used elsewhere in the readout system to receive the fiber inputs for readout. A relative of the CIC card, the analog Serial Fanout (SFO) selectably fans out 1 to 6 UTP Hotlink signals to 12 outputs. This card has several uses in the system, as described below. Finally, the Magic Bus Transceiver (MBT) accepts up to 7 Hotlink inputs, assembles them into events, and broadcasts them on the MBus backplane. The VME backplane is reserved for readout to L3 of events accepted by L2.

The data processing is performed on custom processors, either a serial processor based on the Compaq Alpha chip, or a parallel processor based on TI DSP chips. Alpha boards are easier to program (static-memory C++ under a real-time variant of Linux). The sixteen Alphas are used as crate Administrators, or as Workers for carrying out physics algorithms. The sixteen DSP boards (Second Level Input Computer, or SLICs) are harder to program, and are used for the geometrically-parallel task of handling the muon system track-segment finding. This is the heaviest processing job in L2, and also has the largest number of inputs.

The MBT cards are relatively complex, because they provide much of the system's connectivity and see many different I/O protocols. In addition to event assembly and broadcast, they distribute control information from the L1 trigger framework, do MBus to Hotlink output for the preprocessor Alphas, and send the trigger decision bits to the L2 trigger framework.

The L2 system as a whole occupies 6 VME crates. Most crates can be expanded by adding more input cards and serial processors. The serial processors all see the same input data from the backplane broadcast, so can subdivide the processing task. The SLICs receive their inputs on point-to-point Hotlink inputs, so it is somewhat harder to subdivide their processing tasks.

The Alpha processor board is based on the Compaq PC164 design: a 500 MHz 64-bit 21164 Alpha processor uses the 64-bit 267MB/s PCI bus as a bridge to

commercial Ethernet and VME interfaces and custom MBus I/O, including DMA from the MBT. The card was designed by the CDF group at the University of Michigan, with input from DØ. DØ uses the card for both L2 preprocessors and the L2 global processor, while the rival CDF experiment uses it only for their Global processor.

The SLIC card contains 4 "worker" and one master DSP, each a TI TMS320-C6201. The card contains FIFOs and sophisticated data routing to select which of the 16 inputs are sent to which DSPs. The cards are programmed in C. The SLIC does not connect to the MBus backplane. SLIC outputs are sent on Hotlinks to MBTs which collect track fragments for an Alpha processor to assemble. Thus, the muon processing is actually a 3-stage stochastic pipeline, since the preprocessor itself has two stages. An SFO card is used to fan out L1 control information to each SLIC of a crate, avoiding a separate interface to the L1 framework on the SLIC card.

All cards but the analog SFO and CIC cards have VME interfaces for initialization, and for collection of data for performance monitoring and fault diagnosis.

This system contains many programmed processors. Because taking any element offline impedes normal operations, we built a test facility loosely coupled to the main data taking system. SFO cards make copies of the data entering MBT's or (one of the) SLICs. This allows non-interfering testing of algorithms on live data at full rate before committing the experiment to the new algorithms.

One year after the start of data taking, an upgrade will add a Silicon Tracking Trigger preprocessor to enable triggering on displaced vertices. This project requires 6 crates of electronics, the same as the rest of the L2 trigger. It will send Hotlink outputs to the MBT of the existing tracking preprocessor, supplanting the lower-resolution track information provided by the L1 tracker. VTM modules receive the silicon raw data and the L1 seed tracks. The Fiber Road Card (FRC) controller, the Silicon Track Clustering (STC), and the Track Fit Card (TFC) are all implemented as mezzanines on a common motherboard. Buffering I/O cardlets handle 33 MHz LVDS interconnetions. The FRC and STC mezzanines are implemented by FPGAs, while the TFC uses TI DSPs of a newer and faster generation than the SLICs.

1. R. Hirosky, in *Proceedings of the 11th IEEE Real Time Conference 99*, Santa Fe, New Mexico, 14-18 June 1999, and references therein. See also the Level 2 web site at http://www-d0.fnal.gov ; follow the links Technical, DØ Upgrade, Trigger Systems.