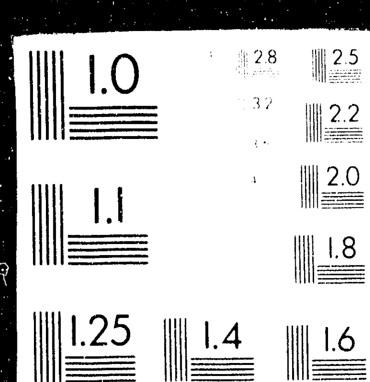
1 OF 1



Conf-

Fermi National Accelerator Laboratory

FNAL/C--92/303

DE93 004743

FERMILAB-Conf-92/303

DØ Triggering and Data Acquisition

Bruce Gibbard
for the DØ Collaboration

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

DEC 1992

October 1992

Presented at *XXVI International Conference on High Energy Physics*,
Dallas, Texas, August 6-12, 1992

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CHO3000 with the United States Department of Energy

Disclaimer

D0 Note 1496 OG 1194
Presented at XXVI Int'l. Conf. on HEP, Dallas, TX, 8/6-12/92
FERMILAB-CONF-92/303

D0 TRIGGERING AND DATA ACQUISITION*

Bruce Gibbard
for the DØ Collaboration†

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

Abstract

The trigger for D0 is a multi-tier system. Within the 3.5 μsec bunch crossing interval, custom electronics select interesting event candidates based on electromagnetic and hadronic energy deposits in the calorimeter and on indications of tracks in the muon system. Subsequent hardware decisions use refined calculations of electron and muon characteristics. The highest level trigger occurs in one element of a farm of microprocessors, where fully developed algorithms for electrons, muons, jets, or missing E_t are executed. This highest level trigger also provides the assembly of the event into its final data structure. Performance of this trigger and data acquisition system in collider operation is described.

INTRODUCTION

D-Zero is a large general purpose detector which began operation at the Fermilab Tevatron in the spring of 1992. It includes 7 major detector subsystems and approximately 115,000 channels of electronics.[1] The detector's run time system consists of a multi-tier trigger integrated with a high performance data acquisition system. The original specification was that it be capable of digitizing and collecting 300 KByte events at rates of 200 to 400 Hz. These events are subjected to high level software filters capable of reducing the rate to approximately 2 Hz which are then recorded. The system is capable of operating in multi-user mode during calibration and debugging in addition to monolithic operation during Physics data taking. In this paper the system will be briefly described and its status

*This work was supported in part by the U.S. Department of Energy and the National Science Foundation.

†The D0 collaboration includes: Universidad de los Andes (Colombia), University of Arizona, Brookhaven National Laboratory, Brown University, University of California at Riverside, CBPF (Brazil), CINVESTAV (Mexico), Columbia University, Delhi University (India), Fermilab, Florida State University, University of Hawaii, University of Illinois at Chicago, Indiana University, Iowa State University, Lawrence Berkeley Laboratory, University of Maryland, University of Michigan, Michigan State University, Moscow State University (Russia), New York University, Northeastern University, Northern Illinois University, Northwestern University, University of Notre Dame, Panjab University (India), IHEP-Protvino (Russia), Purdue University, Rice University, University of Rochester, CEN-Saclay (France), SUNY at Stony Brook, Superconducting Supercollider Laboratory, Tata Institute of Fundamental Research (India), University of Texas at Arlington, Texas A & M University.

and actual observed performance, as of August 1, 1992, reported.

HARDWARE TRIGGER

In Figure 1 are shown the trigger framework and the muon and calorimeter Level 1 and 1.5 systems.[2] Their connection to the digitizing systems and the Level 2 data collection system is also indicated.

Level 0

The sequence of steps in the acquisition of an event begins with the Level 0 trigger which uses scintillation counters to determine if one or more beam-beam interaction occurred during a particular beam crossing. In instances where exactly one beam-beam interaction occurred, a coarse determination of the z position of the interaction vertex is also made.

Level 1

Level 1 decisions are made in the 3.5 μsec between bunch crossings. They are logical combinations of a number of the 256 available "trigger terms". These trigger terms include information from the Level 0 trigger and a variety of physics signatures derived from the other detectors. They include counts of EM towers about E_t thresholds, counts of jet towers above E_t thresholds, missing E_t above thresholds, total scalar E_t above thresholds and the presences of coarse track-like hit distributions in the muon drift chambers.

Level 1.5

For some kinds of Level 1 triggers a Level 1.5 trigger verification is required. Level 1.5 triggers require variable decision time during which dead time is accrued. While Level 1.5 is a generic hook into the D-Zero trigger system, at this time it is utilized only for decisions based on finer grained processing of the muon drift chamber data. The dead time incurred for these decision is typically 5 to 20 μsec. At this time the muon Level 1.5 trigger is being operated in a passive mode which allows determination that it is properly responding to events while not actually determining which events are selected. It is expected to be fully operational within a couple of weeks.

Trigger Framework

The trigger framework which coordinates Level 1 and Level 1.5 triggers is capable of supporting 32 independent triggers, as described above. The trigger framework deals with the digitizing system as 32 independent geographic sectors. It is capable of using any of the 32 Level 1 triggers to initiate readout of any combination of these 32 geographic sectors. Once a Level 1 trigger has been satisfied, including Level 1.5 where necessary, start digitization signals are sent by the trigger framework to the front end electronics of the specified geographic sectors and the Level 2 supervisor which is responsible for data collection is informed of the readout decision. This system is currently fully operational.

DIGITIZING SYSTEM

The VME based digitizing system is housed in 80 crates, each assigned to one of the 32 geographic sectors described above. Flash ADC's are used to digitize the central tracking system and the towers comprising the calorimeter Level 1 trigger. Time multiplexed ADC's are used to digitize the muon system and to do the primary digitization of the calorimeter. All of the digitizing systems have a 4-stage event pipeline in the front end electronics. The digitizing time is specified to be less than 2 msec. At this time the lack of complete list building hardware in the trigger crate has limited the effective digitizing time to 5 msec. This hardware will be augmented

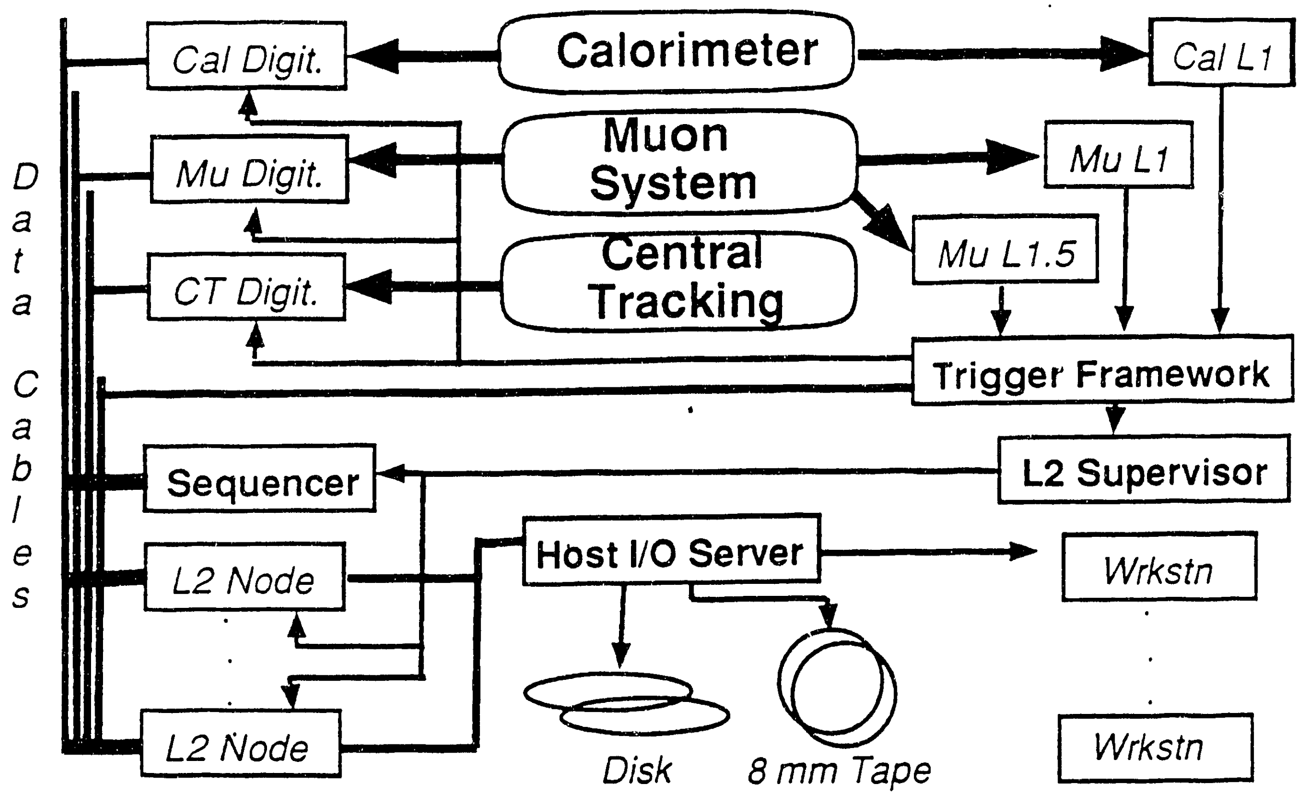


Figure 1. Schematic of the D0 data taking system.

over the next two months. In any case the actual dead time associated with an event under normal operating conditions is approximately 18 μsec, the time required to clear the first state of the pipeline so that a second event can be accepted.

DATA COLLECTION

The data collection system, as indicated in Figure 1 consists of a set of high speed data cables, the VME crate resident cards which drive them, the multiport memory which accepts data from them, the sequencer and the Level 2 supervisor which coordinates their use.[3]

Crate Level Data Gathering

Each digitizing crate contains within it a VME Buffer Driver (VBD) card which contains two buffers and a list processing engine. This engine performs data collection local to that crate. When construction of the local data block is complete (it is being done in parallel in all crates of the experiment), the buffer in which it was gathered is available to be read out onto a custom high speed data cable while the other buffer is being filled from VME with data from the next event. The experiment contains 8 data cables each 4 Bytes wide operated at 10MHz giving a total bandwidth of 320 MBytes/sec. On average there are 10 crates per cable but the actual assignment is done to approximately balance the data load and to isolate detector subsystems from each other. Each cable has a sequencer on it which generates and circulates read-out tokens to the VBD's on that cable thus sequencing their readout onto that cable for that event.

Detectorwide Data Gathering

The data cables are used to move the data from the front end digitizing crates to the farm of VAX processors in whose memory the events are assembled and where high level software trigger decisions are made. Eight channels of multiport memory, one for each data cable, are housed in VME and accessible via a bus adapter by a VAXstation 4000 model 60 processor. The multiport memory is directly mapped into the memory space of the VAXstation so that transparent manipulation of these data is possible. In addition an alternate port to this memory exists which allows direct access to the data by a digital signal processor or other special function processor. Use of this port has not yet been implemented. The data from the cables goes directly into appropriately mapped memories and the necessary structure is put in place around to produce an event record of the type supported by the Zebra memory management software available from CERN. From this point on, all handling of the data is done based on this Zebra structure. While this system was specified as being able to collect 300 KByte events in processors of this farm at rates in excess of 200 Hz, to date actual operation has been at 50 Hz. This rate has been determined by a variety of factors including collider luminosity, recording rate capability, and the need for additional supervisor code to guarantee that events do not sit too long on capacitors in the digitizing systems awaiting readout of earlier events to complete. Remedies to all of these limitations are expected within the next month.

The management of the farm of VAX processors is done by another VAX called the supervisor. Since the code resident in various processors of the farm may be different, routing of particular events to particular processors is based on which of the 32 triggers caused the event to be taken. The supervisor, aware of which processors are qualified to process which trigger types, performs this routing function. Different modes of operation exist relating to whether a single event can be sent to multiple processors or not and whether the system will wait when processors capable of dealing with a particular event type are all busy. The relatively exotic heterogeneous modes of operating the processor farm are used primarily during calibration and debugging. During standard data taking the farm is usually operated in a near homogeneous form. The processor farm consists of 50 nodes each of approximately 12 MIPS yielding a total compute capability of 600 MIPS. At this time only 28 of the 50 nodes are physically installed with the addition of the others expected to match demand as the luminosity grows.

LEVEL 2 FILTERING

Once the event is fully assembled in the address space of one of the processors of the farm, software criteria can be applied to further determine if the event is of interest. The code used to make these software decisions is written in FORTRAN and runs under Digital Equipment Corporation's VAX/ELN operating system. This decision-making is organized around a set of "Tools". When such a tool is invoked it typically calculates some quantities, makes a comparison and then returns either a pass or fail result for that event. Tools associated with the identification and measurement of Jets, Muons, Electrons, Photons, Total Scalar E_t, Missing E_t, Vertex Locations, and Two particle Invariant Masses. have been developed. The complete description of a Level 2 "filter" consists of a script of tools each with its parameters to be applied successively until one of the tools reports failure or the filter is passed. The system can work with up to 128 such filter scripts at a time. The decision as to which scripts should be applied to an event is determined by which Level 1 trigger bits it satisfied. At the completion of the Level 2 filtering phase a 128 bit mask of the filters passed is generated. If the event has any set bits in this filter mask, it will be retained and passed on to the host computer for recording. While rejection rates in the Level 2 filtering system are intended to be 100 to 1, at this time rejection rates of 30 to 1 are being obtained. Additional testing of the filter algorithms on real data is required before the criteria can be safely tightened to give the additional rejection.

DATA RECORDING

The transfer of events from processors in the farm to the host is done by a custom link similar to the data cable by which data is transported from VME crates to the processor farm. The data logging task on the host directs events to various recording streams on the basis of which filter bits have been set. The data of a particular stream are written to disk until a 140 MByte file has been completed. The file is then spooled to an 8 mm tape corresponding to that stream. By using the disk as a buffer, it is possible to record on multiple 8 mm tape drives while retaining event and file order on each tape. There are a total of 8 double density tape drives. The recording system was specified as being capable of recording 300 KByte events at 2 Hz. The actual event size has now been observed to be 450 KBytes per event. Thus though the recording system is operating at 650 KBytes/sec, it results in a rate of 1.3 Hz rather than the specified 2 Hz. The current limitation is the use of an Ethernet link in place of a custom data cable connecting the Level 2 filter nodes to the host. The higher speed link to be installed within the next couple of months will support rates of at least 3 Hz of these larger than expected events.

Via the data logging task events are also made available to tasks running in workstations where online monitoring can be done. The data logging task also directs approximately 10 per cent of the events, selected on the basis of the filter mask, to a small farm of workstations, called the express line, capable of doing the complete reconstruction of such an event sample in real time.

OPERATION

Control of the D0 data taking must of necessity span the various real time subsystems, configuring them for coordinated activity.

Coordinating Task

A single task running in one of the machines of the host VAXcluster performs this function. It runs as a detached server to which clients establish connections in order to request configurational and operational changes in the detector. The coordinating task itself establishes connections to the various run time subsystems to send messages requesting operational changes and the loading of new configurational parameters. This task maintains a model of the run time system which contains the values of configurational parameters as well as the intrinsic and established connectivity across subsystems. Comparisons of change requests with this model are used to determine what allocations should be made and what message sequences are required to reconfigure and/or alter operation of the detector.

Multi-User Operations

Support for multi-user operation of this system was an important original specification. This mode of operation is primarily of use during calibration, testing and debugging operations. The detector most usually operates monolithically for physics data taking. Multi-user operation consists of independent simultaneous runs including separate triggers, filters and recording streams. Allowing for the shared or exclusive use of detector elements facilitates flexible and efficient use of the detector during such operations. Conflict resolution and minimizing the number of configuration messages in this running mode are the responsibility of the coordinating task. Multiuser operation of the detector (typically 4 to 6 users) has been routine for the last couple of years during the assembly and commissioning of the detector.

Specialized Use of Filter Processors

Two specialized uses of the Level 2 filtering system are supported. First, the system can be used to calibrate the detector electronics by performing calibrating calculations in parallel in the filter processors. This allows one to discard the raw data and send only final calibration constants up to the host system to be recorded. In this way one greatly reduces the time involved in doing the calculation and very substantially reduces the load of data which must be transferred to the host computer, thereby reducing, by nearly an order of magnitude, in the time to do a calibration. Second, portions of the the Level 2 filter system can be run in so called "shadow" mode where specified nodes are given copies of a sample of the data which is passing through the system. These nodes then run algorithms currently under test without affecting the actual data selection and recording process. This allows for the real environment testing of new filtering algorithms prior to their inclusion as part of the actual data selection process.

Configuration Descriptions

The D-Zero detector system is configured by a set of files which describe the various triggers, filters, and recording streams as well as the way in which the digitizing system should be configured and read out. These same configuration files can also be used to set up an offline simulation which includes both the hardware trigger (Levels 1 and 1.5) and the Level 2 filter system. In this way Monte Carlo events or previously recorded real events can be subjected to the various selection criteria offline both to certify algorithms and to anticipate their impact on data taking when applied. This same system allows one to calculate efficiencies and rejection rates for the triggering and filtering criteria used to take particular data sets.

REFERENCES

1. See other papers in these proceedings for more detailed descriptions of various subsystems of the D0 detector.
2. M. Abolins, D. Edmunds, P. Laurens, J. Linnemann, B. Pi, "A High Luminosity Trigger Design For The Tevatron Collider Experiment in D0," *IEEE Transactions on Nuclear Science*, Vol. 37, No. 1, 1989, pp. 384–389.
3. D. Cullen-Vidal, D. Cutts, J. Hoftun, D. Nesci, C. Johnson, R. Zeller, "D0 Level-2/Data Acquisition; The New Generation," presented at the Conference on Computing in High Energy Physics, Tsukuba, Japan (March, 1991).

END

DATE
FILMED
2/7/93