SLAC–PUB–5142
November 1989
(E/I)

ISSUES FOR TRIGGER PROCESSING AT HIGH LUMINOSITY COLLIDERS

A. J. Lankford

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309

Abstract

A number of issues for the design of trigger processors at future high-luminosity, high-energy colliders such as the Superconducting Super Collider and the Large Hadron Collider are discussed.

Introduction

Trigger processing is perhaps the most exciting technical challenge at future colliders. It is crucial for extracting the physics signals which we seek to study from extremely high rates of complex background events. In fact, unprecedented interaction rates will require the full power of offline physics analysis techniques to be available in the trigger for event filtering. Consequently, the trigger interacts broadly with both physics goals and detector design.

This workshop contribution identifies some of the issues important to trigger design. It is far from being a comprehensive study. Hopefully it is a provocative introduction to some of the physics requirements and to the range of technical solutions.

Overview of the Trigger

The trigger selects event candidates in a series of stages, or levels, which are progressively more complex and more time-consuming. Each level, by reducing the rate of event candidates, affords the subsequent level more processing time. Although other numbers of levels are possible, this overview will discuss a model trigger with three levels for triggering for high-P_T physics.

Invited talk presented at the ECFA Study Week on Instrumentation Technology for High Luminosity Hadron Colliders, Barcelona, Spain, September 14–21, 1989.

---

*Work supported by the Department of Energy, contract DE–AC03–76SF00515.

At future colliders, even the first stage of trigger decision cannot be made during the interval between bunch crossings. Consequently, every detector signal from every bunch crossing must be buffered until the level 1 trigger decision is complete, and the level 1 trigger must complete a trigger decision each 16 nsec in order to keep pace with the rate of bunch crossings. The level 1 processing time must be minimized in order to reduce the number of bunch crossings for which data will be buffered. Decision times of about 1 μsec are generally discussed in light of the propagation times to and from the trigger on a large detector (about 1/2 μsec) and the need to form some global event quantities such as missing E_T. A fully pipelined hardware processor which exploits extensive parallelism in order to reduce latency will address these requirements. Its pipelined architecture suggests that this processor will have a fixed decision time, which is also convenient for the architecture of the signal buffers. A subset of all detector signals will be provided to the level 1 processor on data paths which are separate from the paths used for data acquisition. The level 1 trigger will provide rejections of between 10^3 and 10^4.

Between 10^4 and 10^5 event candidates per second remain at the input to the level 2 trigger, affording it 10–100 μsec on average per decision. Thus, its processing must be prompt; however, the additional decision time available allows iterative processing, such as sequential processing of track candidates. Additional time also allows event candidates to be directed to independent processors for processing in parallel. In this way, the level 2 trigger can exploit "event parallelism" in the processor farm sense, as well as "parallelism within an event" as used by level 1. With or without the use of event parallelism, microprocessors embedded within the level 2 architecture may play a significant role in the level 2 trigger selection. The level 2 processor will still operate only on a subset of all detector data transported on a separate data path, including the data used by level 1 and the output of level 1.

The iterative nature of level 2 suggests that its decision time will be variable, in the range of tens of microseconds; however, for the convenience of the architecture of the front-end signal buffering, the level 2 trigger processor will preserve the order of event candidates, performing resequencing if trigger decisions complete out of order. Rejections of about 10^2 are expected for level 2.

The rate of event candidates into the level 3 trigger is then between 10^2 and 10^3, a rate which is sufficiently low to allow transport of data from all parts of the detector and to accommodate a farm of microprocessors as the level 3 trigger processor. In fact, rates into level 3 higher than 10^4 may be feasible. The full event, with the full detector resolution, consequently is available, as are the power and flexibility of general-purpose, high-level language programmable CPU's. Rejections of between 10 and 10^2 are expected from level 3, resulting in a final rate of event candidates of a few 10's per second.

Physics Goals

Triggers at future colliders must be designed to identify, count, and measure the quanta which characterize the physics at high energies: jets, muons, electrons, photons, and weakly interacting particles, such as neutrinos, which leave missing E_T. The trigger must also be able to combine requirements on these quanta and on event topology in order to select event candidates. Although triggering on the physics at the highest mass scales will not be difficult, a number of reasons for triggering on physics at lower energies also exist. These reasons include the goal of linking the physics at the highest energies to the physics at present colliders, the goal of studying a relatively low mass (150 GeV) top quark, and the need for adequate calibration events such as W's and Z's. Preliminary studies suggest that the physics goals can be met by prompt triggers which also provide the desired reduction in trigger rates. For instance, for inclusive triggers, thresholds may be set at approximately 40 GeV for inclusive electrons and muons, 1 TeV for single jets, and 175 GeV for missing E_T.

Single Electron Trigger: An Example of a Prompt Trigger

A prompt inclusive electron trigger studied by Sakai of KEK illustrates the nature of selection criteria which might be used and the rate reduction possible. He used a simple calorimetric model with fast shower simulation of QCD events by ISAJET. By requiring a calorimeter tower of size $\Delta\phi \times \Delta\eta = 0.2 \times 0.15$ with electromagnetic energy deposit greater than 20 GeV but with the energy in the hadronic section less than 20% of the energy in the electromagnetic section, he achieved a rejection of greater than 10^4. By also requiring a stiff track ($P_T > 5$ GeV) pointing towards the trigger cell in ϕ (i.e., with no z requirement) and requiring that the trigger cell be isolated (i.e., the energy in nearest neighbor cells is less than 20% than in the trigger cell), the rejection is greater than 10^6 for all energies greater than about 12 GeV.

Although this study deals only with a simplified model of a calorimeter, it suggests strongly that rejections of greater than 10^5 can be achieved by prompt triggers for isolated electrons with P_T greater than 40 GeV.

Inputs to Prompt Triggers

Only a modest fraction of all detector signals is required for prompt triggers. Electron triggers require electromagnetic calorimeter towers of about $\Delta\phi \times \Delta\eta = 0.2 \times 0.2$ over about five units of rapidity, hadronic towers over the same region, and track segments from chambers immediately in front of the calorimeter. By requiring track segments at the outer radius of the tracking volume which point towards the interaction vertex, only stiff tracks (with $P_T > 5$–10 GeV) will be matched to electron candidates in the calorimeter.

Muon triggers will require track segments from the muon chambers, signals from muon system scintillators (if they exist), and track segments from the outer tracking volume. Signals from the hadronic compartment of the calorimeter may also be used.

Jet triggers, and ΣE_T and missing E_T triggers, will require only calorimeter towers which sum the electromagnetic and hadronic portions.

Some General Technical Design Considerations

The bandwidth required to transport data to prompt trigger processors for 60 MHz bunch crossings is quite high, even for subsets of the detector data. For instance, 5000 calorimeter sums of two bytes each require a bandwidth of 600 Gbytes per second .

Most trigger quantities are topologically localized on the detector. For instance, the detector signals which characterize an electron originate in a small region of solid angle. Consequently, much trigger processing could be done locally, which would ease the data bandwidth problem.

Power dissipation of trigger processors, and of drivers which transmit data to the trigger, may limit the amount of trigger processing on various parts of the detector, or it may limit the amount of data which is available to the trigger. For instance, transmission of all hit wire information from a central drift chamber to a remote trigger processor may be problematic, as may be local processing of all hit wires into track segments.

The trigger latency, even for deadtimeless triggers, is important in that it affects the design of front-end electronics. In the simplest solutions, it affects the amount of buffering, and possibly the architecture of the buffers, in the front-end. In some solutions, such as "smart" pixels, the effect on occupancies, ambiguities, and resets is profound. The level 1 latency is at least half a microsecond, which is the propagation time of signals to and from a central trigger processor.

Detector response times and propagation delays within the detector are often longer than the time between crossings. Consequently, signal collection for the trigger, as well as strobes back to the detector, must be time synchronized. Delays will need to be adjusted for groups of channels. Empty beam buckets may help select these delays.

When designing a fast trigger, the designer often has a choice between exploiting event parallelism or parallelism within an event. Event parallelism is exploited by processors working in parallel on separate events, as in a microcomputer farm; whereas, parallelism within an event is exploited by parallel processors working on separate portions, such as different regions of solid angle, of the same event.

The questions of: "How selective should the trigger be?" and "How many events should be written to tape?" are closely related to physics goals. However, there exist tradeoffs between recorded event size and number of events recorded, as well as in applying processing power to reducing one or the other. Both reductions are forms of data filtering.

Event Pileup

Event pileup affects detectors with fast response times, as well as slow detectors, because of multiple interactions per crossing. For an average of 1.6 interactions per crossing, the probability of having more than one interaction is 48%. Given that there was at least one interaction, the probability of having more than one interaction in the same crossing is 60%. Of course, the effect of pileup is smallest for detectors with single crossing response times.

Each detector entity which provides a trigger, e.g., each calorimetric trigger tower, must identify the bunch crossing being triggered upon. Positive crossing identification is possible even for detector components which do not have single crossing response times. For instance, the time of arrival of liquid ionization calorimeter signals can be derived from the zero-crossing of their predictable pulse shape. Time resolution in the 1–2 nsec range should be achievable for 10 GeV electrons and 50–100 GeV jets in liquid argon calorimeters. In drift chambers, correlations in drift times between nearby, offset layers allow untangling of the drift time from the time origin of the ionization.

Event overlap arising from multiple interactions during the resolving time of the detector does not seriously confuse physics. This fact is because the probability of two rare events overlapping to fake a more rare event is small. In addition, the probability of an ordinary event overlapping a rare event to fake a more rare event is less likely than an extra hard gluon radiation within the rare event.

Event overlap does not significantly increase trigger rates for hard processes because it is unlikely to combine hard scatterings from multiple events. Increasing the number of interactions within the resolving time of the detector increases the trigger rate by the same factor; however, it does not change the ratio of accepted to rejected interactions. Isolation cuts, on the other hand, may be compromised by the addition of soft particles within the isolation cone.

ΣE_T is not a good event selection variable because it does not select only hard scattering. Consequently, event pileup significantly increases rates for ΣE_T triggers. Missing E_T, however, is not seriously affected by event pileup because overlapped events do not have large E_T, and hence do not have large missing E_T.

Calorimeter Triggers

Calorimeter triggers require minimal pattern recognition and are naturally implemented in prompt triggers. In fact, the full granularity of the calorimeter, which is driven by detailed e/π separation and by mass resolution for decays of W's into jets, is not needed by the prompt trigger. Consequently, the first step in forming a prompt calorimeter trigger is to sum nearby transverse calorimeter sections into trigger towers with $\Delta\phi \times \Delta\eta$ between 0.1×0.1 and 0.2×0.2. The input signals to the tower sum will be analog, with digitization occurring subsequently; however, particular care must be paid to maintain uniform calibration and timing in order to preserve resolution.

The level 1 trigger can be implemented as a pipelined digital processor, of which the Zeus and D0 level 1 triggers are examples. Digital processing affords well-defined synchronization to a system clock and facilitates, via memory look-up techniques, application of thresholds, weights, and calibration. It will be important, however, to determine the dynamic range which must be maintained during digitization and digital processing.

A variety of prompt jet algorithms are now in use. These include energy clustering about a seed tower as done by CDF, energy summing in overlapped fixed cones as done by UA1, energy clustering in detector subregions with special treatment of edge effects as done by Zeus, and identifying a seed tower only as done by D0.

In order to avoid a separate trigger bias, the trigger should achieve the required level of rejection using the same jet algorithm, or a subset of it, as is used offline for physics analysis. For ease of theoretical interpretation, most experiments now seem to prefer a jet algorithm which defines a jet as energy flow within a fixed cone about a jet axis. The cone size, however, varies with the physics being studied.

What is the ideal prompt calorimeter trigger? Perhaps it would be provided by a massively parallel architecture in which a single, simple processor corresponding to each tower investigates the hypothesis that its tower is the center of an energy cluster (for several fixed apertures), with all towers being processed in parallel, and perhaps even employing the full granularity. A second level of logic could arbitrate overlapping clusters. This trigger implements an offline algorithm with the full resolution of the offline analysis. On the other hand, a much less ambitious solution may also provide the required level of rejection without introducing trigger biases.

Any future prompt calorimeter trigger will more fully exploit the segmentation, calibration, and resolution of the calorimeter than in the past. In fact, few selection criteria may remain for use by the higher level trigger. Higher-level triggers may be limited to refinement of electron identification and further selection and combination of criteria which are formed by the prompt logic.

Tracking Triggers

Tracking of charged particles by the trigger is instrumental to selection of electron and muon candidates. For electrons, the presence of a stiff charged track directed towards an electromagnetic shower reduces photon and π^0 backgrounds. In addition, tracking can link information from transition radiation detectors to showers and can provide an E/p check to help reject chance overlap of a charged track with a shower produced by a photon. Identification of track segments, rather than full track reconstruction and momentum measurement, may be sufficient for any of these tasks.

For high-P_T muons at large angles, sufficient rejection will be provided by simply demanding the presence of a penetrating track segment in the muon system which points back to the interaction vertex, where a cut on the angle of the segment in the bend plane provides a P_T cut. At smaller angles, below about 15 degrees, a sharper P_T cut, in the range of 10–15 GeV will be needed. This requires use of drift time information and track reconstruction even at Level 1.

Beauty physics places a premium on track finding by prompt triggers since the transverse momenta of particles from B decay are not sufficiently large for calorimeter triggers. On the other hand, relatively stiff tracks, in the few GeV range, do arise from the B mass and P_T. A prompt trigger which selects events with at least one track with $P_T > 3$ GeV or at least two tracks with $P_T > 2$ GeV may provide an enhancement in B events of about a factor 50. For this purpose, it may be possible to define a track as a segment at the outer radius of the tracking system whose P_T is measured by linking the segment to the interaction vertex.

Novel techniques for recognizing or measuring charged tracks in prompt triggers include the use of associative memory and of data-driven pipelined processors. Associative memories, including custom VLSI applications for high energy physics, implement template matching techniques which can greatly increase the number of patterns searched as compared to simple memory look-up techniques. The CDF level 2 track finder is an example of a pattern matching segment finder which uses similar techniques to identify tracks with $P_T > 3$ GeV. Data-driven pipelined processors, as implemented for Fermilab E690, can implement track finding and fitting which exploit combinations of parallel processing and processing pipelines to create a machine which is economical in its use of hardware and nearly fully efficient in its use of processing. This architecture could also exploit modern ASIC implementations of many hardware functions, or for that matter could allow embedding of programmable processors for certain tasks.

Higher-Level Triggers

Higher-level triggers will require considerable processing power in order to apply sophisticated event selection criteria to the high input rate of event candidates. Considerable flexibility will be required of the trigger processors in order to allow changes in the event selection criteria as physics experience is gained and as physics goals evolve. This flexible processing power will be provided by large "farms" of powerful commercial microcomputers. For example, between 1000 and 5000 future processors might provide an aggregate CPU power of between 10^5 and 10^6 VAX equivalents, or about 10–100 "VAX-seconds" per event candidate. The processors might be implemented with four RISC processors per board using FUTUREBUS+ with data input via a high-speed external bus. Each processor crate might also include a boot node and a shared mass storage device.

Industry has taken an interest in massive parallel computing on a similar scale. More than one firm now discuss 10^3 to 10^4 parallel nodes for scientific computing using loosely-coupled RISC processors and message passing. Perhaps the future will offer a commercial solution for the higher level trigger.

Such a massive application of processors, however, will demand significant development of software system tools. For instance, the operating system must allow management of data flow and of processing, continuous operation during configuration changes, *in situ* debugging of production code, tools for development of new code, and facilities for verification of proper operation. In fact, the farm must provide a comfortable programming environment with operating system tools comparable to that which exists on the popular minicomputers of today.

Trigger Designer's Tool Kit

The trigger designer today has a wide array of new and more advanced tools available for the task of building fast, powerful triggers. At the component level, programmable logic is available with more versatile cells, larger scale integration, and advanced development (programming) tools. Gate arrays are available in a range of technologies, CMOS, BiCMOS, ECL, and GaAs, allowing optimization of speed and power. They now offer between 10^4 and 10^5 "usable" gates per chip, and will offer more in the future. Silicon compilation offers advanced cell libraries and development tools for semicustom designs, and design of full custom VLSI is possible where required.

For the fastest trigger processors memory look-up techniques will continue to be common for simple pattern recognition and fast mathematics. Content addressable memory, either commercial or custom, offers more efficient use of silicon for pattern matching.

Simple arithmetic processor chips, digital signal processors, and RISC processors can be chosen to match a combination of computational speed and programming flexibility to a task. Modern DSP's are programmable in high-level languages and have versatile operating systems. RISC processors are suitable for embedding in special-purpose devices as well as for general-purpose computing.

Processors with special architectures from outside HEP may also find roles as trigger processors. Image processors offer possibilities for pattern recognition, clustering, and similar tasks. Some of our local pattern recognition problems may be simple to map onto neural nets of realizable scale. Alternatively, neural nets may serve as a paradigm for some application of massive parallel processing. Concurrent machines also offer a form of massive, fine-grained parallelism which may match the topology of some of our processing problems.

Special-purpose processors, such as traditional hardwired triggers, and general-purpose microprocessor farms often seem in competition as trigger processors. In fact, both types of processors have roles in the trigger. Special-purpose processors are necessary for speed at the first levels of prompt triggers, and can be designed to be programmable with respect to important parameters. General-purpose processors are required for flexibility at the last level of event selection. Furthermore, the distinction between special-purpose and general-purpose fades as DSP and RISC cores become embedded in custom circuits and as custom coprocessors are attached to general-purpose CPU's. The crucial issues in choosing technologies are: "How much processing power is required?" and "How much flexibility is needed?" Physics goals and detector design will determine the technology requirements.

Concluding Remarks

Although the trigger problems at future high luminosity colliders are challenging, they are tractable. Thresholds in prompt triggers can be chosen to satisfy both physics goals and data acquisition requirements. Event selection can be accomplished online with the same programmable flexibility available for offline physics analysis. Technology for electronics which can meet the processing challenges is rapidly developing.

Acknowledgments

This workshop contribution summarizes many contributions to previous workshops by other authors, who I would like to acknowledge here.