Trigger and Data Acquisition at High-Rate Colliders

A. J. Lankford
University of California
Irvine, CA 92717

Abstract

Trigger and data acquisition systems for experiments at future e^+e^- and hadron colliders are discussed. The challenges and their solutions in the two environments are compared and contrasted. Experiments for the asymmetric B-factory PEP II and the SSC are used as examples.

1. Introduction

Future e^+e^- and hadron colliders pose many similar electronics challenges to experiments. These similarities arise from the experimental need for high luminosities, high frequency bunch crossings, highly granular complex detectors, and sophisticated event selection. The similarities of the challenges give rise to the potential for similar architectural solutions for trigger and data acquisition. Nonetheless, differences in the fundamental physics processes give rise to important differences in the detailed requirements for trigger and data acquisition in the two environments. These lectures survey many of the similarities and differences in the requirements for trigger and data acquisition systems at e^+e^- and hadron colliders, and will discuss possible solutions to the technical challenges. They will draw upon the examples of the asymmetric B-factory PEP II as a future e^+e^- collider and the SSC as a future hadron collider.

Section 2 will present an overview of the trigger and data acquisition problem and of the general architectural solution. Section 3 will outline the general functionality of front-end electronics, and Section 4 will sketch data acquisition systems, with examples from the SSC and PEP II. Section 5 will discuss triggers, again with examples from the SSC and PEP II. Finally, Section 6 summarizes.

These lectures will not provide a general introduction to trigger and data acquistion systems. Such introductions can be found in previous SLAC Summer Institute lectures by Marty Breidenbach[1] for data acquisition systems and, for trigger systems, in the first chapter of a book by Bock, Grote, Notz, and Regler on Analysis Techniques for High-Energy Physics.[2] Many additional review papers and papers on interesting developments can be found in the Computing in High Energy Physics conference series and in the IEEE Real-Time Computer Applications in Nuclear, Particle, and Plasma Physics conference series.

2. Overview and General Architectural Solution

The technical challenges of trigger and data acquisition systems at future colliders stem from the fact that the study of rare physics processes demands high luminosity. In the environment of future colliders, high data rates will arise from a number of factors, particularly, the high luminosity and the large cross sections for competing physics processes. In addition, highly granular detectors with large numbers of channels to dissect complex events and multiple data samples per channel, for instance from waveform sampling electronics, will contribute to high data rates. Backgrounds of accidental hits from high current beams or from large cross section physics processes will also contribute. Coping with these high data rates will require increases in three basic parameters of the trigger and data acquisition systems: increased trigger sophistication, increased data acquisition bandwidth, and increased online processing power.

This overview chapter starts with a comparison of parameters for future expertments at SSC and at the asymmetric B-factory PEP II. This comparison will serve to illustrate the challenges to trigger and data acquisition systems at future colliders. The following sections outline the new architectural requirements and sketch a basic architectural approach that uses a multilevel trigger and data acquisition system. Finally, the essential characteristics of trigger levels and basic functions of electronics subsystems are discussed.

2.1 Comparison of SSC and PEP II Parameters

To consider some of the similarities and differences between the experimental environments and trigger and data acquisition requirements at future hadron and e^+e^- colliders, examine the comparison of SSC and PEP II parameters presented in Table 1. The design luminosities of both machines are much higher than luminosities at present colliders. To achieve these high luminosities, both machines will collide incident bunches at high frequency. At SSC, the time between bunch crossings will be 16 nsec. At PEP II, the time between crossings will be only 4 nsec. Both times are orders of magnitudes shorter than most existing colliders, and still much shorter than the 96 nsec existing at HERA.

Table 1. Comparison of SSC and PEP II Parameters.

Parameter	SSC	PEP II
Luminosity ($cm^{-2}sec^{-1}$)	10^{33}	10^{34}
Time between Crossings	16 nsec	4 nsec
Inelastic Cross Section	100 mb	5 nb
Total Rate	10^8 Hz	50 Hz
<# Interactions/Crossing>	1.6	10^{-6}
Channel Count	10^6	10^5
Silicon Vertex Detector?	yes	yes
Event Size	1 MByte	25 KBytes
Events to Tape	100 Hz	100 Hz

The total cross sections are of course much different in hadron and e^+e^- colliders. At SSC, the total inelastic cross section will be about 100 mb, which coupled with the design luminosity will give a total interaction rate of 10^8 Hz. This high rate gives 1.6 interactions on average per bunch crossing. At PEP II, the total annihilation cross section, even on the $\Upsilon(4S)$ resonance, is only 5.5 nb. Consequently, the total interaction rate is 50 Hz and the average number of interactions per crossing is only 10^{-6}.

Detectors for experiments at both machines will be highly granular, although SSC detectors (10^6 channels) will be more highly segmented than a detector at PEP II (10^5 channels), in order to cope with higher particle multiplicities. In both cases, the experiments will also have silicon vertex detectors.

The different particle multiplicities in the two environments will result in different sizes for event records. Typical SSC events may be as large as 1 MByte; whereas, typical PEP II events will not be much different in size than events in current e^+e^- detectors, perhaps 25 KBytes.

Typical maximum rates of writing events to tape are likely to be similar in the two environments. In both cases, the rates for physics processes to be recorded are about 50 Hz. The SSC rate is given by the cross sections for W's, Z's, and top quarks decaying to high-p_t leptons, and the PEP II rate is given by the $\Upsilon(4S)$ cross section. An event writing rate of about 100 Hz represents a reasonable goal for the purity of the event sample, *i.e.*, 1:1 signal to noise.

The difference in total rate, or in physics interactions per crossing, is the

greatest difference between the hadron and e^+e^- environments, and gives rise to significantly different trigger strategies, as explained later. Nevertheless, the data acquisition requirements, arising from high bunch-crossing frequency and relatively large rates of events to tape, are sufficiently similar that the overall trigger and data acquisition architectures are likely to be similar in the two environments.

2.2 New Architectural Requirements

As already explained, high luminosity is accomplished with high frequency beam collisions. At PEP II, the beams will collide at 250 MHz. At SSC, they will collide at 60 MHz. In addition to the problems of extremely high rates and very large numbers of channels, certain challenges arise from the short times between bunch crossings which result from the high frequency collisions. In particular, the time between bunch crossings will be less than typical detector response times. It will also be less than the time of flight within the detector, and it will be significantly less than the trigger decision time.

The high data rates and short time between possible interactions of interest will require systems with new features. Firstly, the electronics systems must be virtually without deadtime. They must recognize and collect detector signals at the same time as processing and transporting large amounts of digital data. Secondly, the trigger and data acquisition systems must be pipelined. Pipelined trigger systems must conclude trigger decisions with the same frequency as the crossings. Pipelined data acquisition systems must buffer data from multiple crossings while the trigger is processing.

2.3 Basic Architectural Approach

A basic architectural approach which one can adopt in order to address these new trigger and data acquisition challenges starts with choosing a coherent architecture for the readout of all detector subsystems. Of course, this coherent architecture must meet the requirements of all the individual subsystems, which is also a challenge. A single architecture, which achieves as much commonality as possible among subsystems, is required to optimize cost, reliability, and ease of debugging. These three desirable features are necessities in a large electronics system.

Another aspect of the basic architectural approach is to perform as much signal processing on the detector as is practical, particularly data filtering and multiplexing prior to readout. This approach minimizes the bandwidth of data transmitted from the detector, hence reducing system cost and complexity. Although performing processing of data on the detector, if it involves buffering of data on the detector, entails separate dedicated data paths to the trigger, the benefits in cost and complexity generally outweigh the costs of the dedicated paths. Processing of the data on the detector is made practical by extensive use of custom integrated circuits to accomplish high channel densities and low cost.

The third major aspect of this architectural approach is to exploit parallelism throughout the architecture to achieve high performance. The use of parallelism will avoid bandwidth bottlenecks and allows scaling to the level of performance required, providing the capability to upgrade performance in response to changes in physics goals or luminosity.

The last principal feature of this architectural approach is to implement a multilevel trigger and readout architecture. Such an architecture makes efficient use of bandwidth and processing. Figure 1 sketches the data flow through a multilevel architecture with three levels. The trigger selects events in a series of progressively more complex, more time-consuming stages, or "levels." By reducing the event rate, each level reduces the data bandwidth and affords the subsequent level more time for its decision. In Fig. 1, the full event data is shown moving down the readout chain, on the left side of the figure, from the detector through three levels (L1 Buffer, L2 Buffer, and L3 Farm) and then to mass storage. Trigger data is moved from the detector to the Level 1 trigger (L1 Trigger) in order to be used for event selection. Data bandwidth is reduced by using only a subset of the full data. Following the Level 1 trigger, its output "accepts" or "rejects" an event candidate. For accepted candidates, the partial data used by Level 1 is augmented by additional partial event data and processed by the Level 2 trigger (L2 Trigger). Again, data bandwidth is optimized by only moving the data required by the Level 2 trigger. Finally, for event candidates which have been accepted by Level 2, the full event data is passed to the Level 3 farm, which does the final event selection and outputs accepted event candidates to mass storage.

It is important to emphasize that the choice in Fig. 1 of three levels, rather than more or fewer levels, is arbitrary. The appropriate number of levels depends on the details of the detector, its physics, and its trigger strategy. The algorithms which select event candidates at each level of the trigger will determine both the data bandwidth required for input into the trigger processors and the data rate between stages of the data acquisition. Within a given experiment, a certain amount of flexibility will be available with respect to choosing at which trigger level to deploy selection criteria; however, the algorithms are detector and physics dependent. For instance, at the SSC, the trigger criteria and the final rate of interesting events are quite different for experiments studying high-p_t phenomena and ones studying decays of beauty. Thus, a high degree of interplay exists between the capabilities of the trigger and of the data acquisition at each level in the system. To a large extent, the overall architecture of the detector readout systems will be determined by the architecture of the

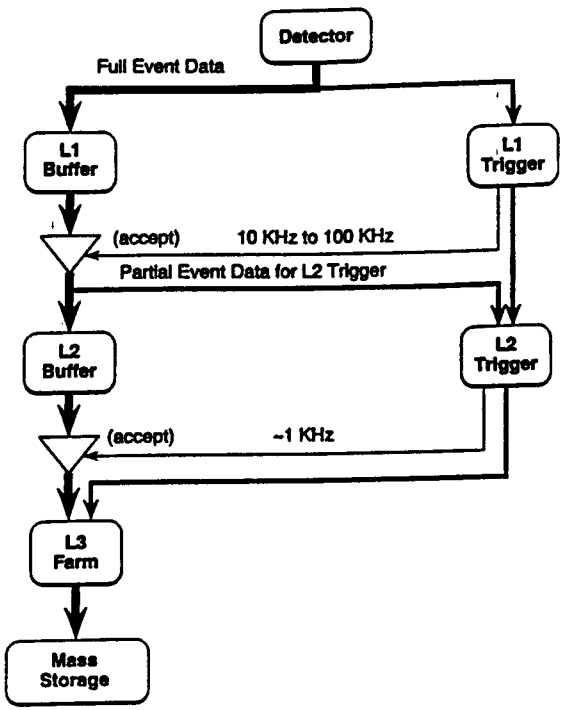

Fig. 1: Data flow diagram of a multilevel trigger and data acqusition architecture with three levels.

trigger. Figure 2 shows the example of a three-level architecture applied to a SSC detector.

2.4 Characteristics of Trigger Levels

Figure 2 and Table 2 illustrate parameters of each trigger level for the example of a SSC detector.

Table 2. SDC Trigger Levels and Rates.

Parameter	Level 1	Level 2	Level 3
Average Time between Decisions	16 nsec	10-100 μsec	1 msec
Average Decision Time	3 μsec	50 μsec	0.5 sec
Input Rate (Hz)	10^8	10^4-10^5	$\sim 10^3$
Expected Rejection	10^3-10^4	10-10^2	10-10^2
Output Rate (Hz)	10^4-10^5	$\sim 10^3$	10-10^2

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 (at the SSC) 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 2 to 4 μsec are generally discussed in light of the propagation times to and from the trigger on a large detector (about 1 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 to 100 μsec on average per decision. Thus, its processing must be prompt; nonetheless, 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

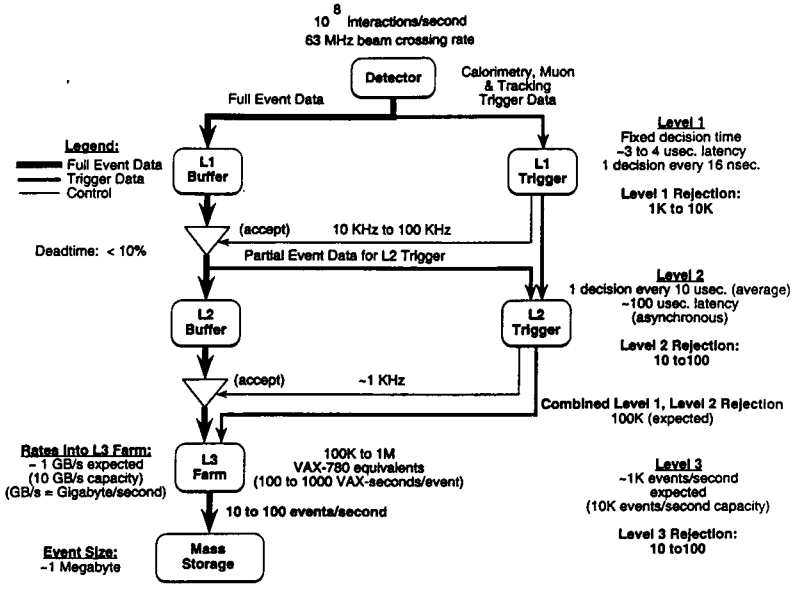

Fig. 2: Data flow diagram of a three-level trigger and data acquisition architecture applied to the SDC Experiment for the SSC (from Ref. 3).

"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. Level 2 is likely to provide a transition from the dedicated hardware processor of Level 1 to the general-purpose CPU of Level 3 by using elements of both. 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 to 100 are expected for Level 2, achieving a combined rejection for Levels 1 and 2 of about 10^5.

The rate of event candidates into the Level 3 trigger is then about 10^3 Hz, 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 Hz may be feasible, and the example shown in Fig. 2 provides a capacity of 10^4 events per second. The full event, with the full detector resolution, consequently is available, as are the power and flexibility of general-purpose, high-level language programmable CPUs. Rejections of between 10 and 100 are expected from Level 3, resulting in a final rate of event candidates of a few 10s per second.

2.5 Functions of Electronic Subsystems

Figure 3 shows a block diagram of the electronics subsystems, from the particle detectors to archival storage. The three principal subsystems are: the front-end electronics, the data acquisition system, and the trigger system. This figure is from the Technical Design Report of the SDC Collaboration.[3]

The front-end electronics processes detector signals, correlates them with particular beam crossings, buffers them during trigger decisions, filters them in accordance with Level 1 and 2 triggers, digitizes the event data so selected, and outputs "event data fragments." The front-end electronics also processes data into "trigger primitives" for the trigger system.

The data acquisition system collects event data fragments from the front-end electronics and assembles the event fragments into complete event records. It also filters the event records according to Level 3 trigger algorithms and records selected events for offline analysis.

The trigger system receives the trigger primitives from the front-end electronics, processes the trigger data, and selects event candidates for further pro-

cessing. The trigger system also provides the control of the front-end electronics, the data acquisition system, and the trigger system.

3. Front-End Electronics

The term "front-end electronics" is broadly used to denote the electronics that is physically located within the detector. The basic functionality of the front-end electronics was outlined in Section 2.5. Its basic purpose is, for those events selected by the trigger system, to convert sets of signals from detector elements into digital event fragments.

This chapter first describes the architecture of front-end electronics, using a block diagram of a channel of a "generic" front-end. It then outlines the motivation for developing custom, detector-mounted, integrated front-end electronics. Finally, two approaches to calorimeter readout are presented as examples of possible front-end electronics implementation.

3.1 Architecture of the Front-End Electronics

A conceptual diagram showing the required functionality for a "generic" set of front-end electronics is shown in Fig. 4. As drawn, the figure shows a single channel that processes the signals from a single detector element. In practice, large numbers of parallel channels synchronously perform identical processing functions on arrays of detector elements. The electronics of all detector components are expected to have similar architectures, enabling a common control and readout scheme for the entire detector. Note that the functional architecture shown in Fig. 4 is the "logical" architecture, not necessarily the physical architecture, of the front-end electronics. Figure 4, and its discussion are adopted from Section 8.1.4 of the Technical Design Report of the Solenoidal Detector Collaboration.[3]

"Signal processing" typically consists of low-noise, high time resolution amplifiers and signal shapers. Each amplifier receives a signal, almost always an analog signal, from one detector element, such as a silicon strip, a wire, or a photomultiplier tube. The signal is shaped to extract the best quality time and/or amplitude information from the signal. Analog-to-digital conversion may occur immediately after this stage, or it may occur later in the chain.

The provision for generating calibration signals via a "calibrator" is typically provided. This feature can be used for channel time and/or amplitude calibration by inserting artificial sets of data to test all or part of the electronics chain. The "calibration strobe" signal controls the timing of the calibration signal.

Whether digital or analog, signals must be preserved within the front-end electronics for the time required to make a Level 1 trigger decision. In the "Level

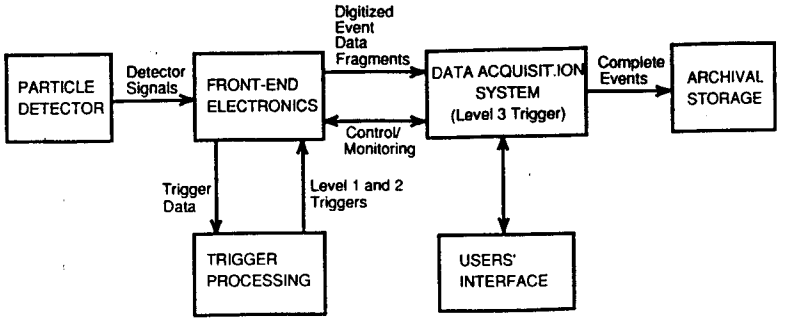

Fig. 3: Simplified block diagram of electronics subsystems (from Ref. 3).

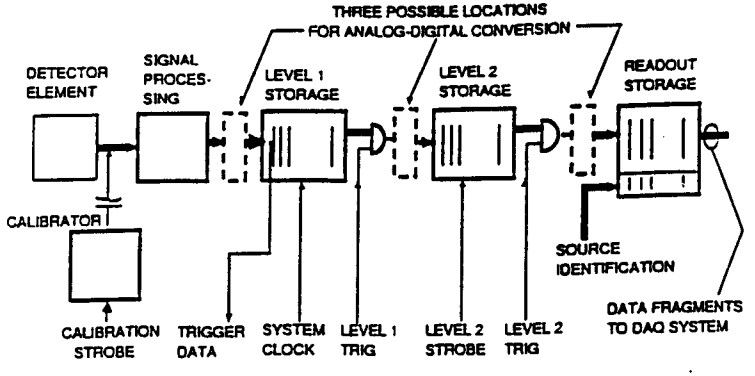

Fig. 4: Block diagram of a channel of "generic" front-end electronics, showing the logical architecture (from Ref. 3).

1 storage," the correlation of stored data with the beam crossing of origin must be preserved so that the proper data is selected by a "Level 1 trigger" decision to accept the data, and so that the data can be correlated with data stored in other parts of the detector. The "system clock," which cycles each beam crossing, indexes the data in the Level 1 storage.

The very small fraction of data that are selected by Level 1 triggers is placed in the "Level 2 storage." The Level 2 storage time, typically 10 to 100 μsec, is longer than that of Level 1. Nonetheless, the volume of stored data will be much less because of the Level 1 trigger selection. In some systems the transfer of data from Level 1 to Level 2 storage is physical, and in others only pointers are transferred. When analog data is stored, usually only pointers are transferred. The time of transfer of data out of Level 1 storage is another possible moment to perform analog-to-digital conversion.

A Level 2 accept or reject decision is received each time the Level 2 trigger system completes its processing. The "Level 2 strobe" signals the completion of the Level 2 decision. A "Level 2 trigger" accept causes the Level 2 storage to output the next event data which it stores. Outputting data from the Level 2 storage may involve an analog-to-digital conversion, if the data has not already been digitized.

As the data is output from Level 2 storage, it is usually merged with other physically local data and buffered in "readout storage," queued for readout by the data acquisition system. The event data are tagged with salient "source identification" which might include identification of physical origin of the data, beam crossing number, and other items of use for later processing. This local data is one of many event "data fragments" that the data acquisition combines to assemble a complete event record.

3.1.1 Analog-to-Digital Conversion

Data fragments delivered to the data acquistion system must be digital. Therefore, an analog-to-digital conversion must be located somewhere between the detector element and the data acquisition system. Figure 4 shows three possible locations.

Analog-to-digital conversion at the earliest opportunity, prior to Level 1 storage, must be performed at the high rate of 60 MHz. The Level 1 and Level 2 storage is then easily done digitally. This option is the choice made for most tracking systems which perform simple discrimination for hit identification, perform low dynamic range digitization, or involve time measurement.

For conversion after a Level 1 trigger accept decision, analog storage must be provided for the duration of the Level 1 trigger latency time. Digitization then takes place at the Level 1 accept rate, and data storage is digital pending a Level 2 trigger decision. This mode of operation may be useful in cases where digital detector signals are not required for output to the Level 1 trigger processor but are required for the Level 2 trigger processor.

Conversion after a Level 2 trigger acceptance will be at the lowest possible rate, between 1 and 10 kHz; however, in that case, analog storage must be provided for both the Level 1 and Level 2 latency times. This mode is frequently the choice for large dynamic range systems, such as calorimeter readout.

3.1.2 Nature of Storage

Figure 4 shows two separate storage entities for Level 1 and Level 2. In fact, both storages can be accomplished within a single storage structure, such as switched capacitor array. There, by means of address manipulation, a single location in an array of storage elements can be assigned first as "Level 1 storage," then "Level 2 storage," followed by "readout storage," and finally "available for storage." In systems using analog storage, this scheme avoids the need to rapidly transfer analog data from one storage location to another.

3.2 Detector-Mounted, Integrated Front-End Electronics

Many data acquisition functions which traditionally have been executed in the counting house, such as digitization, multiplexing, and buffering, will be performed in highly integrated, detector-mounted electronics at future colliders. For some time now amplifying electronics have been mounted in close proximity to particle detectors for improved analog performance and for increased immunity to RF pickup. More recently, silicon microstrip detectors and the SLD collaboration have pioneered detector-mounted custom VLSI and hybrid circuits. In the case of silicon microstrips, the density of connections and limited space for cables leading to and from amplifiers, particularly in four-π detectors, have led to amplifiers, sample-and-holds, multiplexers, and in some cases sparse-scanned readout on a single custom VLSI chip. The SLD collaboration has incorporated similar functionality into all of their electronic systems, through the use of hybrid and semi-custom monolithic amplifiers and custom sampling and buffering ICs. SLD's goal was to achieve a more cost-effective, space-efficient, and reliable electronics system for a detector with a large number of channels.

The trend towards detector-mounted custom ICs must be carried further in the future. Previous motivations, including improved analog performance, increased immunity to RF pickup, density of connections, limited cable space, cost effectiveness, space efficiency, and reliability, are joined by the compelling needs of reduced power dissipation and of increased functionality (*e.g.*, multiple event buffering, integrated trigger solutions, and simultaneous readin and readout). Multiple event buffering is needed to buffer signals from each of the many beam-crossings (128-256 at the SSC) which occur during the time required by the

Level 1 trigger decision. This buffering is performed in the front-end electronics in order to eliminate the power dissipation required to drive the signals off the detector. Integrating trigger processing into the front-ends also reduces the amount of data to be transmitted to the trigger from the front ends, thus reducing power dissipation and interconnections. Simultaneous readin and readout, that is continuing to sample subsequent crossings at the same time as buffered data is read out for triggered interactions, is necessitated by the desire to limit deadtime in face of the high interaction rate.

Consequently, the front-end electronics for future collider experiments typically will be implemented in a pair of multichannel custom integrated circuits. The usual arrangement will consist of a bipolar IC that provides signal processing and a CMOS IC that provides data storage. For future experiments, these chips will replace both the boxes of detector-mounted amplifiers and the crates of remote FASTBUS or VME modules found in today's large detectors, as well as the hundreds of long cable interconnections. Further detector-mounted multiplexing and data preprocessing will replace today's crate-level scanners and segment interconnects.

3.3 Calorimeter Front-End Electronics

Calorimeter front-end electronics currently under development can be considered as examples of future front-end electronics systems. The requirements upon front-end electronics for calorimetry are more challenging than for other detector subsystems. The calorimeter front-end electronics must measure pulse heights with very high dynamic range. It must sample the calorimeter signals at the bunch-crossing frequency. It must also have the capability to read out multiple samples per event, or to combine (with weights) samples nearby in time. Finally, it must perform analog signal processing simultaneous with output of digital data, without deadtime. The dynamic range requirement makes the other requirements more severe, particularly the sampling requirement and the simultaneous readin/readout requirement.

The dynamic range requirement is set by minimum and maximum detectable signals. The SDC Experiment has set the minimum detectable signal at $E_t = 20$ MeV. This value is chosen to be sensitive to 100 MeV of transverse leakage of an electromagnetic shower into adjacent calorimeter towers. This value is also sensitive to the 250 MeV energy deposit of minimum ionizing particles. SDC has chosen to set the full-scale signal at $E_t = 4$ TeV. At this value, saturation is extremely rare for processes that have measurable cross sections. For instance, dijets with $m_{jj} > 20$ TeV deposit $E_t > 4$ TeV in the electromagnetic calorimeter in only three events per year. These minumum and maximum values set a dynamic range requirement of 2×10^5, or $\sim$18 bits. For comparison, the trigger dynamic range is 4×10^3, or 12 bits; furthermore, for the trigger, a nonlinear 8-bit digitization is sufficient.

In order to address the design challenge of readout with high dynamic range at high sampling frequency, at least two different approaches are being pursued. In the first approach, which is called digital PMT readout, the calorimeter signal is digitized at the beam crossing frequency with a "floating-point" flash encoder. The second approach stores an analog signal in an analog memory, or Switched Capacitor Array (SCA), until digitization after a Level 2 trigger.

The calorimeter readout with analog memories provides 18-bit dynamic range with two 13-bit scales. This bilinear scheme has been used often for systems with about 15-bit dynamic range. The conceptual architecture of this approach is illustrated in Fig. 5. The preamplifier/shaper is located near the photomultiplier (PMT). The preamplifier provides the full dynamic range, which is possible in a bipolar technology. The shaper uses delay-line shaping and provides a dual-range output. The rest of the system is contained on a readout card which provides Level 1 and Level 2 analog storage in two channels of SCA, digitizes after an accept decision from the Level 2 trigger, and outputs data to the trigger and to the data acquisition. An address list processor controls the read and write addresses of the SCAs, performing address manipulation as sketched in Section 3.1.2. In this approach, the dynamic range can be limited by cell-to-cell variations in the SCAs. To date, groups at LBL and elsewhere have been successful at designs which control these variations, and which avoid cell-to-cell calibration corrections.

The digital PMT readout provides dynamic range of 18 to 20 bits using ten scales of 8 to 10 bits each. It digitizes at the beam crossing rate, enabling it to provide the trigger with digitizations of full resolution. The digitizer circuit is mounted in the PMT base. It is connected by a ribbon cable to a crate-based digital readout card which contains a calibration lookup table, Level 1 and Level 2 storage, and outputs to the trigger and data acquisition. In this approach, the large dynamic range is restricted to the first element of the system. This element is a custom splitter/integrator IC. Its first stage splits the PMT current into ten binary ranges. The second stage integrates the current in each range onto a storage capacitor (one in a round-robin set of four) during a 16-nsec gate. The third stage consists of a comparator, latch, and encoder which select which range is of interest. The final stage is an analog multiplexer which switches the storage capacitor for the range of interest into the input of a nearby flash ADC. A block diagram of the splitter/integrator IC is shown in Fig. 6. The output of the flash ADC provides an 8 to 10 bit mantissa, and the encoder provides a 4-bit exponent, for a 12-bit pseudo-floating-point output from the base. The splitter must have the full dynamic range; however, the integrator, multiplexer, and ADC are only required to have 8 to 10 bit dynamic range. A Fermilab

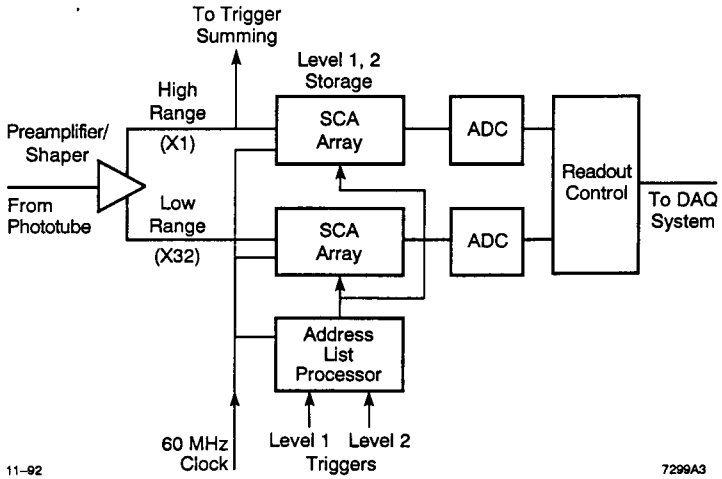

Fig. 5: Conceptual architecture of a channel of calorimeter front-end electronics based on analog data storage in switched capacitor arrays (SCAs). Two SCAs are used to cover a larger dynamic range (from Ref. 3).

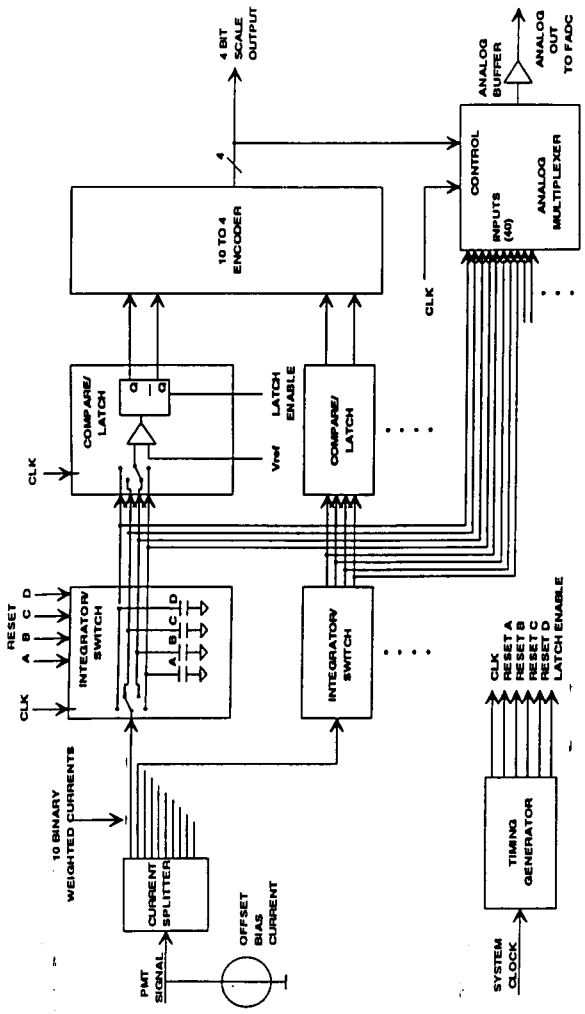

Fig. 6: Block diagram of a splitter/integrator IC for readout of calorimeter photomultiplier tubes of the SDC Experiment. The output of the splitter/integrator is connected to a Flash ADC which digitizes at the bunch crossing frequency of 60 MHz (from Ref. 3).

group has designed splitters and gated integrators with the requisite dynamic range. The SDC Experiment has chosen to use this approach for its calorimeter readout.

The two approaches described share many basic similarities. In particular, they can interface to the calorimeter and to the trigger and data acquisition systems in similar ways. The approaches differ in the number of ranges which they employ to achieve the full dynamic range. The readout with analog memories uses only two ranges, in order to minimize the number of components and connections. The digital PMT readout uses ten ranges, in order to digitize at the bunch-crossing frequency. The approaches also differ in the physical location of the digitization. The readout with analog memories digitizes at a location removed from the input of the preamplifier, in order to minimize digital-to-analog crosstalk within each channel. The digital PMT readout digitizes in each phototube base, in order to minimize the crosstalk betweeen channels. Other differences between the two systems arise from these basic differences. Each approach offers its attractions and disadvantages; however, both approaches are likely to yield solutions which provide the requisite dynamic range in full system implementation.

4. Data Acquisition Systems

As outlined in Section 2, the data acquisition system must transport up to 10 GBytes per second from the front-end electronics to Level 3, it must provide the processing power for the Level 3 trigger, it must control the flow of data from the front ends through Level 3 and to mass storage, it must monitor the operation and the performance of the detector, and it must achieve a manageable, cost-effective solution. In order to achieve these goals, the data acquisition architecture will exploit extensive parallelism in highly buffered data collection from front-end electronics, in its data links, and in event building. The architecture will also be modular and scalable, both in its hardware and software aspects. Data acquisition will also make extensive use of commercial hardware and software from rapidly evolving computer and communications industries.

This chapter first outlines an architecture for future data acquisition systems and discusses the aspects of the architecture. It then briefly presents examples of SSC and PEP II data acquisition systems, using the SDC Experiment as the SSC example. Finally, it discusses some issues important to the design of large electronics systems.

4.1 Data Acquisition Architecture

The principal features of a future data acquisition architecture are shown in Fig. 7. Following an "accept" signal from the Level 1 & 2 trigger system, "data collection electronics" collect the data from the detector-mounted front-end electronics of the various detector components. High-speed data links, using fiber optics, transport data to a "parallel event builder." The parallel event builder organizes event data fragments from various detector elements into complete event records. Processing elements, such as DSPs, can be incorporated at a number of points along the data path. Industry-supported communications links transport data from the event builder to a "parallel processor farm," which performs final event selection.

Computing facilities will also be provided for permanent recording of accepted events, the user interface, and detector monitoring. Simple control mechanisms will keep data flowing at high rates. In fact, separate data and control paths are likely. Networks will be used for system initialization, downloading, calibration, and monitoring. Processors will be extensively used for triggering, calibration, data compression, and monitoring tasks, and their extensive use will lead to an increased dependence on software.

4.1.1 Data Collection

Although each type of detector component will have custom front-end electronics appropriate to its measured quantities, the control and readout of the front-end chip sets will be sufficiently similar that a common readout scheme may be achievable for the entire detector. Data from as many as several hundred thousand front-end chips, each with data rates as high as hundreds of KBytes/s must be multiplexed onto a manageable number, perhaps 100 to 1000, high-speed data channels which provide an aggregate data rate between several GBytes/s and 100 GBytes/s. A hierarchical solution to data collection, starting with groupings of nearby detector channels and proceeding towards large groupings of all the data from one region of solid angle, will be appropriate. The entire data collection process, reducing the number of data paths to the few hundreds to be input to the parallel event builder, is likely to occur within and on the detector.

Figure 8 schematically shows a possible scheme for the first stage of data collection from the front-end chips. A dedicated data collection chip (DCC) is linked to the data outputs of a set of front-end chips (FE #1, ... , FE #n) by a "parallel bus structure." Control of the front-end chips, which could include chip initialization as well as readout control, is shown in this figure as a "serial data bus." At the output of the DCC the combination of parallel bus structure and serial data bus are replicated in order that the subsequent stage of data collection

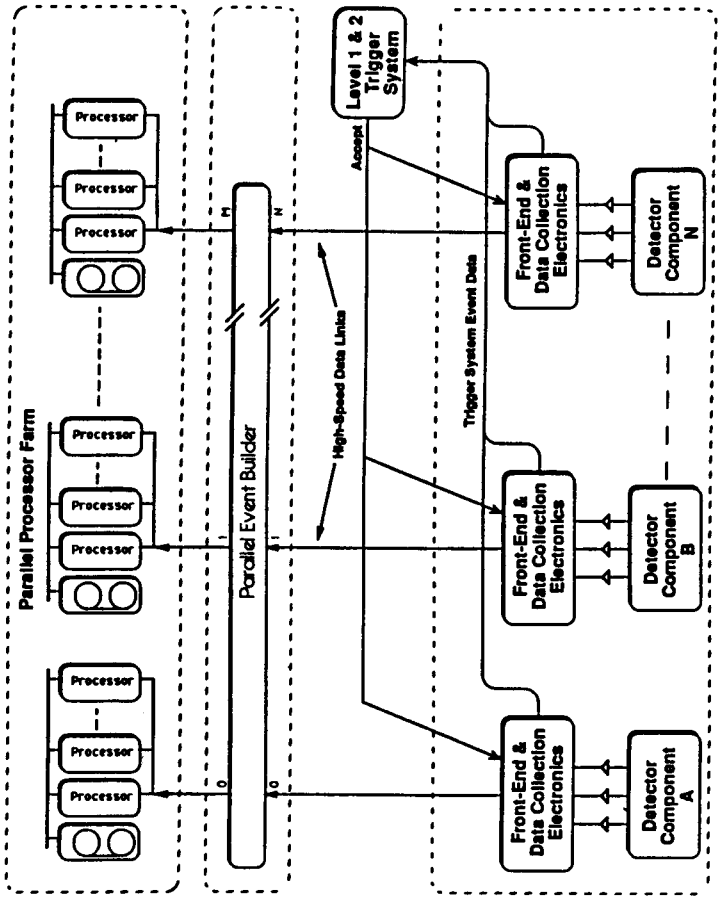

Fig. 7: Simplified block diagram of the architecture of a future data acquisition system (from Ref. 4).

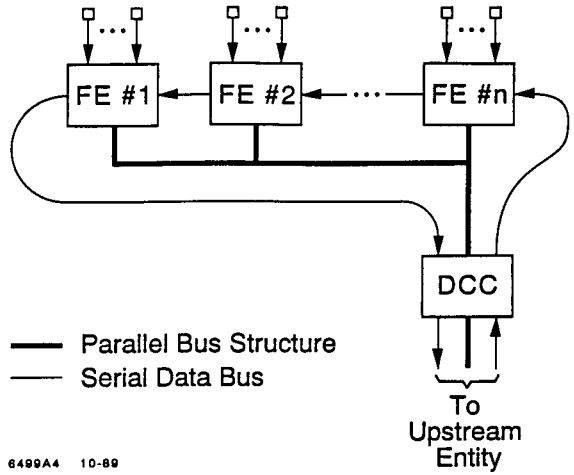

Fig. 8: Scheme for data collection from front-end chips, showing a parallel bus structure for event data readout and a serial data bus for control.

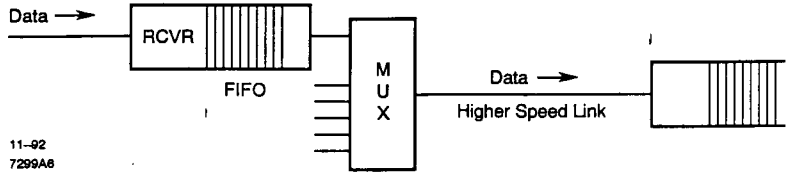

Fig. 9: Block diagram of one stage of a high-speed, data-driven data collection system.

can be done with an identical DCC in a hierarchal manner. This scheme, as presented, is only a rudimentary example of the functionality required. Other data collection architectures are of course possible. For instance, serial data links from each front-end chip to the DCC, connected in a star, might be used to replace the parallel bus structure with a potentially more reliable solution.

The most ambitious solutions to the problem of data collection, those aimed at the highest achievable rates of data transfer, are data driven. At each step in the data collection process, every data source is pushing data into intermediate buffers as the data becomes available. As illustrated in Fig. 9, multiplexors (MUX) then gather the data from the buffers at the highest possible rate and push the data into the next stage of buffers. The bandwidth of all data links can be used to full efficiency. The data is transmitted with appropriate event and channel tags; however, packets of data do not necessarily correspond to individual events. The process of event building is therefore to a large extent decoupled from the data collection and transmission. In these data-driven schemes, control is minimized as data is moved along a series of simple data transmission links. Control occurs on paths separate from the data paths. Operation of such a system should be easy to verify and trouble-shoot, since verification and fault identification will be amenable to a series of communications tests.

4.1.2 Data Transmission

Transmission of data to each stage of data collection will be via links of technology appropriate to the bandwidth required at that stage. Data collected from the front-end chips, where bandwidths are low, will be transported via copper busses on detector-mounted printed circuit boards. At the other end of the data collection process, the perhaps hundreds of long links carrying the data from all parts of the detector to the parallel event builder in the control room will be high-speed fiberoptic links. The speed and number of links at that stage will be determined by practical considerations, such as the cost and the size of the switching network in the event builder. The transition from high-speed copper links to fiberoptic links of modest speed will occur at some intermediate stage.

The principal advantage offered by fiberoptic transmission is high bandwidth, particularly over distances longer than several meters. Fiber optics promise performance that makes data acquisition of GigaBytes per second feasible. Fiberoptic transmission also offers the important advantages of immunity to electromagnetic interference and low transmission losses. In addition, if used within the detector, they offer advantages in size and mass over copper cables. Radiation hard fibers are available to a level of some MegaRads and exhibit some self-annealing. The logic necessary to drive and receive high-speed fiber optic systems can also be rad hard to the same and higher levels.

The fiber optic needs of the computer industry are driving technology to increased performance and decreased cost for links similar to those needed for high performance data acquisition. As shown in Fig. 10, a fiberoptic data transmission system consists, at the transmitting end, of logic to encode the data, including a conversion from parallel to serial bit streams, as well as logic to apply the protocol of the system to the data transmitted. At the receiving end, the complementary functions are performed. The data is decoded and converted back to parallel form from serial. In addition, there is protocol and clock recovery logic. Typically, a fiberoptic data transmission system is bidirectional, utilizing a pair of fibers. Much of the required logic for Gigabit per second systems is now available commercially. For instance, Hewlett Packard markets a chip set dubbed the G-link that operates well in excess of a Gigabit per second and is currently priced in the range of hundreds of dollars per set. Prices are expected to drop markedly further, making such links quite accessible for use in future high energy physics experiments. In addition, HEP groups at Oxford and Rutherford Lab are working on the development of very low-cost 60 MHz fiberoptic systems as an economical solution for systems requiring many links without the demand for extremely high performance.

4.1.3 Parallel Event Builder

The parallel event builder addresses the bandwidth bottleneck arising in traditional event builders where data all passes through one path. In a parallel event builder, a number of input data paths from the detector are connected to a number of output data paths to the processors, and all the data paths can be active simultaneously to maintain the aggregate bandwidth. A parallel event builder is shown very schematically in Fig. 11. The number of input and output data paths need not be equal; however, if bandwidth is nearly optimized then the numbers are naturally the same.

Several schemes for parallel event builders have been discussed. These schemes generally utilize a matrix of buffer/router nodes or utilize switching networks. A few schemes are shown in Fig. 12. Note that, in this figure, the block shown as a processor (denoted by a "P") may be a set of processors. In the first scheme, a multi-drop scheme, data sources (denoted by "S's") deliver data to processor memories (denoted by "M's") in parallel. An entire event is built in the set of memories on the processor bus. This is the scheme used by D0. The second scheme employs a buffer/router matrix, utilizing dual-port memories. The data sources deliver data to a column of memories. The processors have access to an entire event in a row in the memory array. The first two schemes are topologically equivalent. The third scheme utilizes a switching network to interconnect data sources and processors. Memories are used to buffer the data as it passes out of the sources and into the processors. The

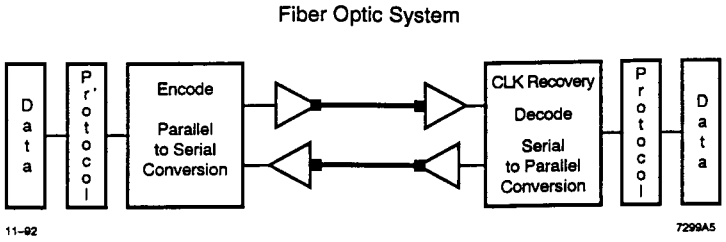

Fig. 10: Simplified block diagram of a fiber optic data transmission system, including logic for transmitting and receiving data.

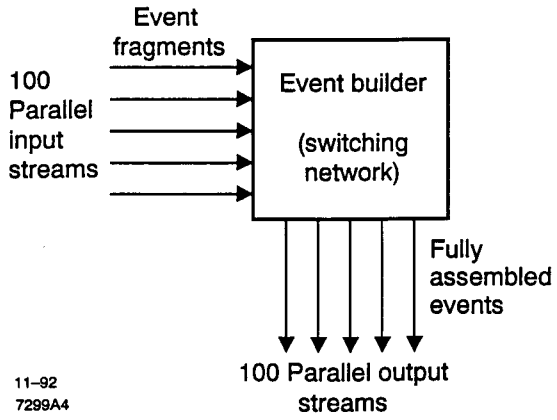

Fig. 11: Illustration of the input and output of a parallel event builder.

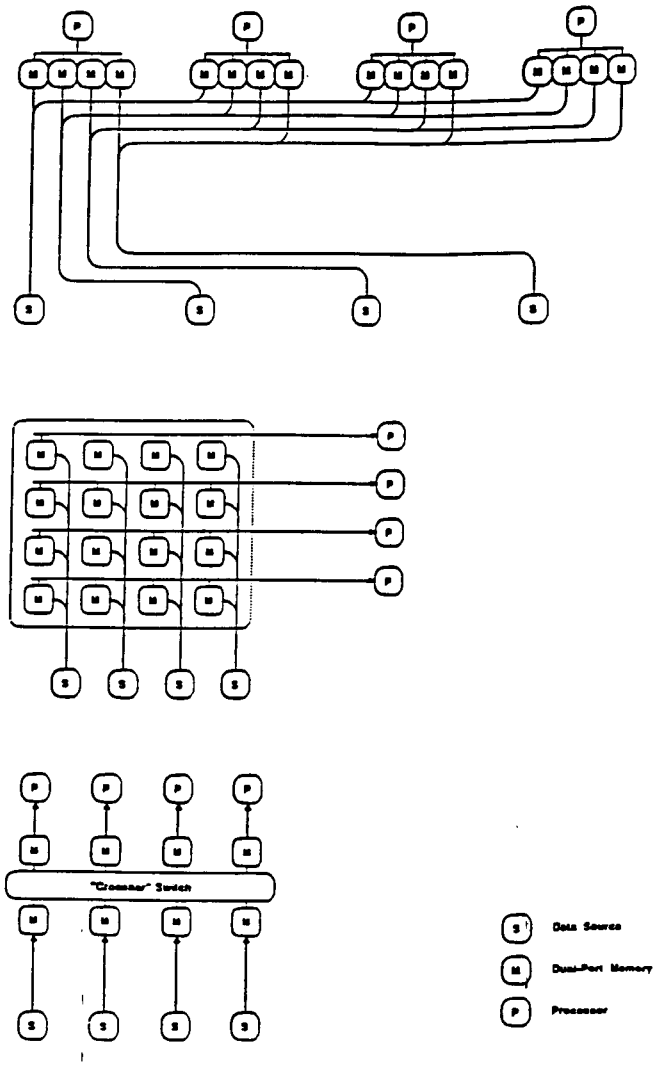

Fig. 12: Illustration of three possible schemes for parallel event building. "S's" denote data sources, "M's" denote dual-port memories, and "P's" denote processors. Scheme A is a multi-drop scheme. Scheme B uses a buffer/router matrix. Scheme C employs a switching network.

CDF Experiment has adopted a switching network, the Ultranet Hub, for event building in a future upgrade. Parallel event builders using switching networks exploit advances in the technology of cross-point switches arising in the communications and computer industry. A generalized $n \times n$ network would allow the interconnection of input and output data paths in any combination; however, events can be built using simpler networks. An interesting implementation for very high-performance systems and which minimizes control is that of the "barrel-shifter" originally suggested by Bowden and Barsotti of Fermilab. In this scheme, each input path is sequentially connected to an output path in a cyclic fashion.

Parallel event building schemes have many similarities, particularly the need for extensive buffering to smooth out event-to-event fluctuations in amounts of data on each link and the need to balance the average data rates on each data path. These needs arise from the fact that the bandwidth will be limited by the longest event fragment of each event if the buffers are insufficient or by the slowest data path if rates are not balanced.

The problem of parallel event building can also be rephrased as one of delivering data from a number of "source nodes" to a number of "destination nodes." The problem then can be addressed by standard computer networking solutions, making available the technological developments of the computer industry.

Any one of these several hardware alternatives could provide the necessary bandwidth for parallel event building, even at the very high data bandwidths which will be required at the SSC. Consequently, the complexity of the software necessary to control the flow of event data fragments through the event builder is likely to distinguish one hardware solution from another. At present, solutions based on commercial switching networks and protocols, such as Fiber Channel and ATM, are generally favored.

4.1.4 Mass Storage

The required bandwidth for recording selected events at future colliders will be in the 10 to 100 MBytes/s range (see Table 1). Very large volumes of recorded data, 100 to 1000 Terabytes per year per experiment, must also be handled.

Although optical tape and disk drives will probably provide the highest bandwidths and data densities in the future, parallel output data streams utilizing more conventional drives and media can provide the required bandwidth. In fact, parallel data streams may also be desired to record different event types on separate drives. Parallel drives can be interfaced to the processor farm through switching networks which tie the drives to the busses on which the processors reside. Helical scan magnetic tape technology developed for the commercial broadcast industry can provide storage media of sufficiently high density (200 GBytes/cassette) for the expected large data samples at the same time as providing drives in the 15 to 30 MByte/s range.

An alternative to directly recording the output event stream at the site of the experiment is to transmit the data via a high-speed link to the site of the offline computing. At the offline site the data can be recorded by a robotic data archiving system which is shared by offline computing. Fiberoptic systems with the necessary bandwidth for the high-speed link now exist and will be commonplace in advance of the operation of future colliders. Standard protocols for such links are now being developed.

4.1.5 Online Processing

Two categories of parallel processing exist in a large collider experiment. A processor farm performs data processing and event selection on data from the entire detector, with each processor executing the same program on a separate event. Other processors, distributed throughout the architecture of the online system, preprocess streams of data from portions of the detector and control and monitor detector components.

4.1.5.1 Online Processor Farm Requirements

The highest level of processing for event selection will generally occur in a farm of many microprocessors which may be characterized by its input and output bandwidths, its processing power, and its software environment.

The required input bandwidth to the farm is dependent upon the physics goals of the experiment and upon the deployment of trigger selection criteria between low-level trigger processors and the online processor farm. The aggregate bandwidths most often discussed for the SSC are between 10 GBytes/s and 100 GBytes/s. The 10 GBytes/s rate arises from a conservatively designed data acquisition system for a high-p_t experiment with a prompt trigger rejection of 10^4, *i.e.*, 10^4 event/s $\times$ 1 MByte/event = 10 GBytes/s. Clearly, an experiment with a prompt trigger rejection of 10^5 to 10^6 would require less input bandwidth. On the other hand, a B-physics experiment operating at a hadron collider with $\mathcal{L} = 10^{32} cm^{-2} sec^{-1}$ with a prompt rejection of about 10^2 would require input bandwidth of 100 GBytes/s. These bandwidths to the parallel processors can be provided by parallel data links. As discussed previously, the required output bandwidth from the farm to mass storage is between 10 and 100 MBytes/s. Parallel output data links can also be used. The aggregate processing power of farms at the SSC is usually described as being between 10^5 and 10^6 MIPs. These estimates are loosely based upon needing approximately 100 to 1000 MIP-seconds per event to perform final event selection.

The architecture of the online processor farm must allow execution of background tasks to the event selection process. Such tasks include testing of new trigger code in parallel with the execution of standard code, verification of event selection processing, and detector performance monitoring. This requirement demands the ability to share events or data among processors.

At least four options exist for the implementation of the farm using commercial products. Commercial microprocessors could be implemented on custom processor boards. Although this approach has been chosen in the past, it no longer seems cost-effective. The other options are: commercial single-board computers implemented as processing nodes, commercial workstations as processing nodes, or a commercial multiprocessor system implemented as the entire farm. Commercial workstations are presently used for most processing farms; however, the last option may be made possible by the growing interest of industry in large-scale application of parallel processing for general scientific computing problems.

An open architecture is another often mentioned requirement of the farm. A truly open architecture would allow one to exploit the most cost-effective microprocessor at the time of system implementation, instead of committing to a processor at the time of system design. This point of view is reinforced by the tendency to employ as much computing power as can be made available, and by the frequent need to expand the available computing power.

4.1.5.2 Distributed Processing and Control

Although the largest scale use of commercial processors in future collider experiments will be in the processor farm, commercial processors will be used extensively for other functions throughout the data acquisition architecture. Processing functions will largely be of the same nature as in current experiments; however, the amount of processing will substantially increase. More than in the past, standard microprocessors will be found embedded in special-purpose low-level trigger processors, in data preprocessors, and in detached control processors. Commercial processors will continue to serve as hosts for the system as a whole and for each detector subsystem. Workstations will be used to interface physicists to the online system, to control and monitor the detector and its performance, and as powerful online graphics machines.

4.1.5.3 Some Software Requirements for Online Processors

The software environment provided for online processing, particularly in the farm, is of critical importance. The farm must execute large programs written for offline processors, which implies that the farm processors must have high quality compilers compatible with those used offline. The farm must provide a code development environment which facilitates production and initial debugging of new code, or be compatible with such an environment on another machine. It must also offer adequate tools for *in situ* debugging of code during operation, *i.e.*, debugging of code executing on any node in a multiprocessor system and debugging of interprocessor communications. Code running on such a powerful machine will require new levels of reliability because of the tremendous number of instructions being executed per second. In addition, the operating system must provide tools for data transfer to and from processors and for control and monitoring of processors. In short, the farm must provide a software environment as comfortable as provided by the previous generation of popular minicomputers.

4.2 Data Acquisition at the SSC: An Example from SDC

As an example of a future data acquisition system for an SSC experiment, consider the system being designed for the SDC Experiment.[3] As the first step in the design of its data acquisition system, SDC has defined a set of requirements. This set consists of requirements for performance, for partitioning and stand-alone operation, for control and monitoring, and for scalability, reliability, and maintainability.

Its performance requirements include:

- maximum Level 3 trigger system input rate of 10 kHz;
- 394 independent data sources;
- maximum bandwidth through event builder subsystem of 10 GBytes/sec (based on 10 kHz and 1 MByte/event);
- minimum processing power in online farm of 10^5 MIPs;
- minimum archival storage rate of 100 MByte/sec;
- maximum event size for a calibration event of 20 MBytes;
- expected event size of 1 MByte;
- 10% maximum readout deadtime; and
- 5% maximum deadtime due to data acquisition errors and downtime.

The principal partitioning and stand-alone operation requirements are:

- must be able to operate separate non-interfering data acquisition systems for each subsystem during commissioning; and

- preserve the above functionality after detector commissioning for debugging and calibration of individual subsystems.

The control and monitoring requirements include:

- set up (download) entire detector into known state;
- track operation of both data acquisition system and detector subsystems;
- record conditions under which data are collected;
- allow for calibration data acquisition;
- allow for non-event data acquisition;
- detect and record error conditions; and
- prioritized alarm system.

A simplified block diagram of the SDC data acquisition system is shown in Fig. 13. The data acquisition system consists of 274 readout crates containing front-end and trigger electronics. These crates provide 394 sources of event data fragments to an event builder subsystem. From the event builder, the data flows to an online computer subsystem, through a computer network, and into online storage. The control of the data flow is provided by the event data flow control subsystem, which coordinates with the Level 1 and 2 trigger systems.

Each readout crate contains front-end or trigger modules, along with a standard set of control modules. The control modules include a commercial data acquisition CPU, a crate adapter/interface module which provides the high-speed link to the event builder, and a trigger gating module. The crates are tied together by a control/monitoring network and by a network of fast trigger cables. In some cases where individual front-end modules output high data rates, data links to the event builder originate directly from the front-end modules. SDC crates are sketched in Fig. 14.

The SDC event builder subsystem consists of a switching network, which is expected to be commercial, and data balancing and input queuing logic, which balances the flow that arrives on the input links before the data is input to the switching network. Among other possible commercial solutions, SDC is considering Fiber Channel as the fabric for the event builder.

Although the switching network architecture is homogeneous, one solution for the control of data flow imposes a flexible, reconfigurable hierarchal architecture of data collection and event building onto the network. In this solution, event data fragments from approximately 20 sets of 20 data sources flow through the switch to event building nodes. These nodes assemble subevents which they output to processors in the Level 3 farm where complete events are assembled. In order to efficiently utilize the bandwidth of the network, the small event data fragments, which are on average 2.5 KBytes, are buffered into messages of about 500 KBytes.

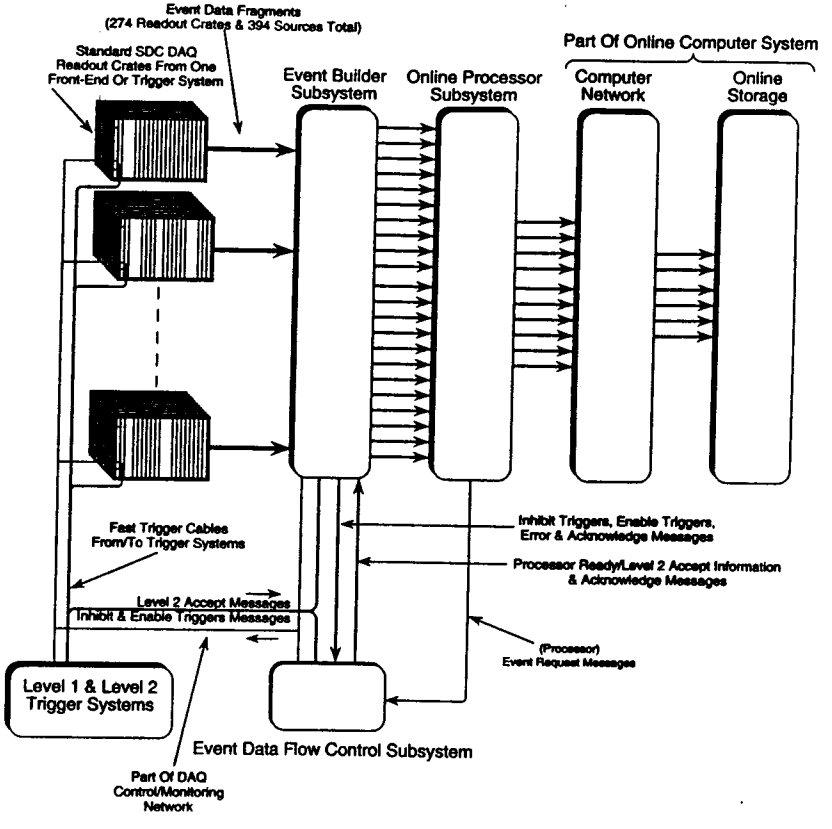

Fig. 13: Simplified block diagram of the data acquisition system of the SDC Experiment (from Ref. 3).

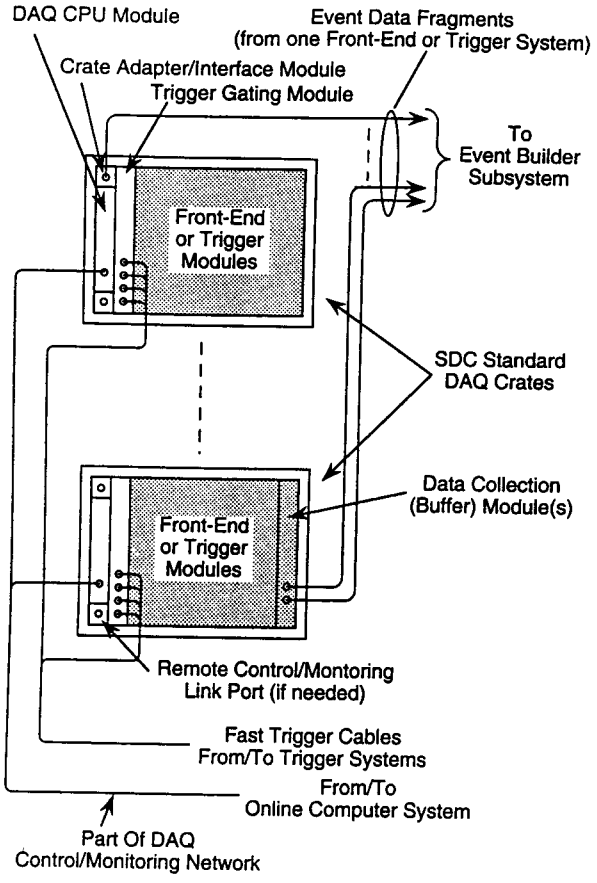

Fig. 14: Illustration of standard SDC crates, including standard control modules and network links (from Ref. 3).

The control/monitoring network ties together all the data acquisition processors, online computing workstations, and the slow controls system. This network is expected to be a commercial successor to Ethernet.

The SDC conceptual design seeks to use commercial hardware and software to the best possible advantage, in order to achieve a cost-effective, high-performance solution which can be adapted to the changing needs of the experiment. The conceptual design is still at an early stage of development, and may change significantly before implementation.

4.3 Data Acquisition at PEP II

An experiment at PEP II will also make heavy use of commercial products. A block diagram of a potential system is shown in Fig. 15. In this system, subsystem computers for the data acquisition and trigger electronics are read out through a buffered switch to event-level computers of the Level 3 trigger. Although performance does not need to be as high as at the SSC, the architecture to be adopted is quite similar. The buffered switch is analogous to the SDC event builder subsystem, with switch and data balancing and input queuing logic. Data flow is controlled via a separate control network as in SDC. The workstation farm is also tied together by a commercial network, which is expected to be fiber optic.

4.4 Some Issues in Design of Large Systems

The systems necessary to address requirements of triggering, data acquisition, and online processing for future collider experiments will be substantially larger and more complex than the corresponding systems in existing experiments. Consequently, new issues arise in the design of these large systems.

Functional modelling (*i.e.*, behavioral simulation) of the overall system, including the trigger, data acquisition, and processing, will be necessary to study system performance with respect to many parameters and to verify system design. The overall system can be modelled at a high level. Mixed level simulation will be needed to simulate components at various levels of detail in the context of the overall system design. Tools for mixed analog and digital simulation of the demanding front-end electronics would be extremely useful.

The overall design must not allow system complexity to scale with the number of detector channels. Readout solutions should be integrated across detector components. Control mechanisms should be simple.

The applicability of commercial developments and of emerging technologies must be monitored for performance and cost advantages. The overall system architecture should permit the exploitation of technical advances which occur during the development of the experiment, and even during its operational phase.

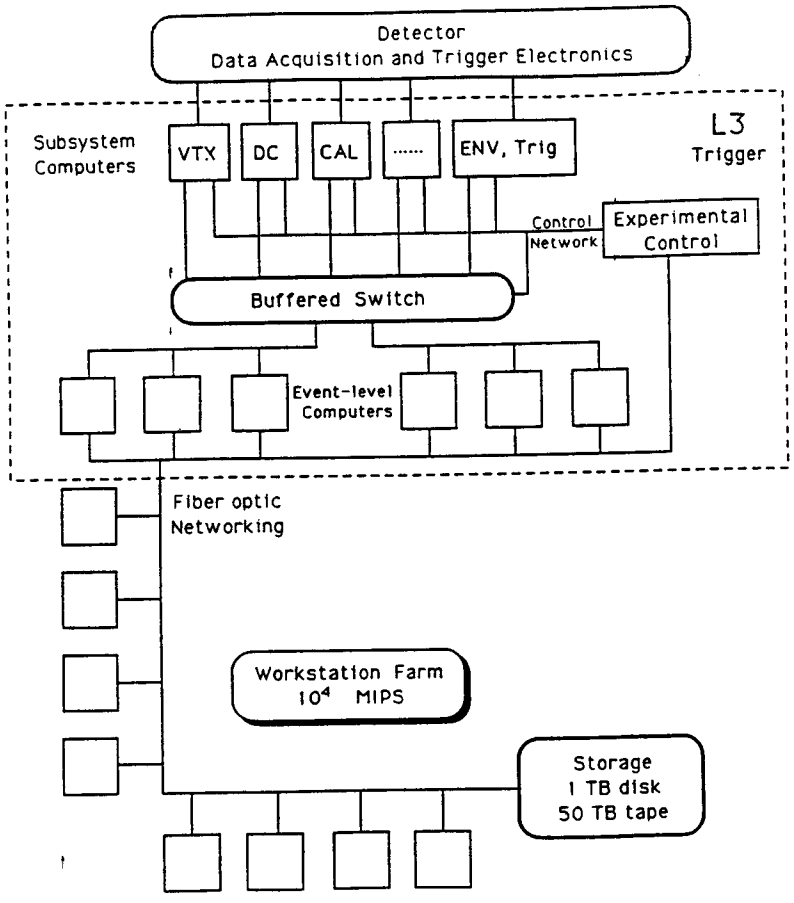

Fig. 15: Simplified block diagram of the data acquisition system of an experiment for PEP II (from Ref. 5).

Issues of reliability, redundancy, and in some cases radiation tolerance will require additional engineering techniques and skills. Finally, the verifiability and maintainabilty of the very large systems must be considered throughout the design.

5. Trigger Systems

The trigger system performs two closely related but distinct functions. The first is to identify particle interactions, *i.e.*, to identify the occurrence of an interaction of an incident particle with a target particle which scatters interaction products into the detector. The second function is to select interactions for which to acquire data. At e^+e^- colliders, the central problem for the trigger is to identify physics interactions, since virtually all interactions are recorded. At hadron colliders, the central problem is to select which physics interactions to record, since nearly every beam crossing produces an interaction. Recall the comparison of SSC and PEP II parameters in Table 1. The rate of bunch crossings are similar, 60 MHz and 250 MHz respectively, and the rate of events to tape is about 100 Hz in both cases. However, the total rate of interactions are widely different, 10^8 Hz at SSC and about 50 Hz at PEP II.

As explained in the introduction, the similar large reductions from the native rate, *i.e.*, the rate of bunch crossings, to the events-to-tape rate in the two environments leads to similar multilevel trigger and data acquisition architectures for the two environments. Recall that the trigger architecture, although largely designed to facilitate data acquisition, determines the data acquisition architecture. Throughout this chapter on trigger systems, recall also the general characteristics of the trigger levels described in Section 2.4.

5.1 Some General Trigger Processing Issues

The bandwidth required to transport data to prompt trigger processors for bunch crossings in the 60-250 MHz range is quite high, even for subsets of the detector data. For instance, 5000 calorimeter sums of 2 bytes each require a bandwidth of 600 GBytes/s at 60 MHz.

Most trigger quantities are topologically localized in the detector. For instance, the detector signals which characterize an electron originate in a small region of solid angle. Consequently, much trigger processing can be done locally, which eases 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 designer will often have 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 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 located on a large detector. The Level 1 latency defined for future experiments generally ranges between two and four microseconds.

Processing must be provided such that each detector entity which provides a trigger, *e.g.*, each calorimetric trigger tower, can 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 to 100 GeV jets in liquid argon calorimeters with signal shaping appropriate to the high rates of the SSC. In drift chambers, correlations in drift times between nearby offset layers allow untangling of the drift time from the time origin of the ionization. Examples of techniques for identifying the bunch crossing are presented in the following Section 5.1.1.

Processing to identify or disentangle multiple interactions during the resolving time of the detector will also be needed.

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, tradeoffs exist 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.

5.1.1 Bunch-crossing Identification

As mentioned above, positive bunch-crossing identification is possible even for detector components which do not have single crossing response times. In wire chambers of certain geometries, the drift time can resolve the bunch-crossing time. As shown in Fig. 16, in a chamber with half-cell offset between layers of drift cells, the sum of the drift times (t_2 and either t_1 or t_3) for drift to the wires on either side of an ionization track is a constant equal to the total drift time (T). In the figure, the average of the two drift times to the right (t_1 and t_3) is used for a three-layer chamber.

As another example, the time of arrival of liquid ionization calorimeter signals can be derived from the zero-crossing of shaped pulses. The principle is illustrated in Fig. 17 by an analog example. Similar digital processing could also be applied for the case where the analog calorimeter trigger signal is already waveform digitized. A "trigger sum," derived by summing the analog signals in a trigger tower, is input to a "threshold unit" which outputs a signal for the time that the input signal is over threshold. The "time-over-threshold" signal is delayed by the signal risetime, so that it is centered on the expected zero-crossing time of the signal. The time-above-threshold signal enables a "zero-crossing unit" which also receives the trigger sum as input. The zero-crossing unit outputs a pulse at the time of the zero crossing of the trigger sum, if the sum is above threshold. The output pulse of the zero-crossing unit enables a "synchronization unit," which synchronizes a delayed copy of the "bunch-crossing" signal, to produce a "Level 1 accept" pulse. The Level 1 accept pulse therefore is only generated if the trigger sum is above threshold, and it is synchronized with the zero-crossing time of the trigger sum and with the bunch-crossing signal.

5.2 Triggers at the SSC

The trigger is the start of the physics event selection process. The event candidates rejected by the trigger can never be recovered for physics analysis. At hadron colliders, this fact is particularly meaningful, because the trigger must select which of many physics interactions to retain for future analysis. At the SSC, the physics challenge is to retain all physics over the wide range of energies and masses from the TeVatron to the 40 TeV of the SSC. The technical challenge is to address the physics challenge while accomplishing a factor of 10^6 rejection.

The solution which addresses these challenges will be a multilevel trigger with nearly the same sophistication as offline physics analysis. It will exploit simple fast electronics at the first levels and high performance commercial processors at high levels, and it will transition from simple to more complex processors at intermediate levels.

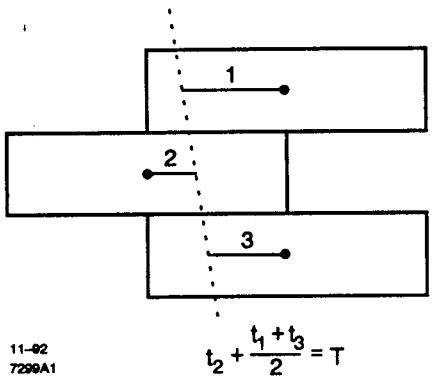

Fig. 16: Technique for deriving the time of traversal of an ionizing track passing through a drift chamber composed of layers with half-cell offset. The sum of the drift times to the left and to the right is a constant equal to the maximum drift time. This technique can be used to identify the bunch crossing from which the track originates.

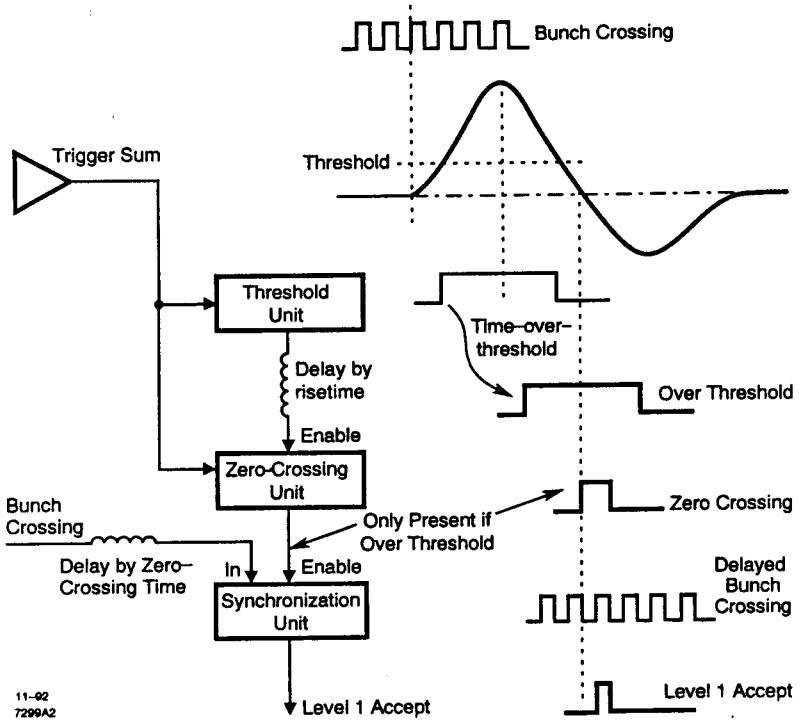

Fig. 17: Technique for identifying the bunch crossing of origin for a liquid ionization calorimeter signal using the zero-crossing time of the signal.

An approach to the design of this complex trigger starts by specifying the physics quanta upon which to trigger and by defining the criteria by which the quanta are identified. Once the criteria are identified, it is then possible to specify the detector data required by the identification criteria. Understanding the criteria and the requirements for detector data, the criteria then can be defined as algorithms and assigned to trigger levels. At this point, the trigger strategy and architecture are sufficiently well defined to allow the design of data paths to the trigger levels and to design trigger processors at each level.

Table 3, from the SDC Technical Design Report,[3] shows the physics signatures of some sample physics processes of interest at the SSC. The physics goals involve signatures comprised of high p_t electrons, muons, photons, and jets and of missing E_t. Consequently, the trigger must identify, measure, and count these quantities. The trigger should identify the basic physics quanta by their local signatures in the detector, with minimal use of topological criteria (*e.g.*, isolation) particularly at the earliest level of the trigger. The trigger must also achieve thresholds of interest within the allowable trigger rates. Moreover, the selection criteria employed by the trigger should be compatible with (and not determine) the identification criteria which will be used for offline analysis, and the trigger should be measurably efficient. At the SSC, signatures of standard physics processes can be used as benchmarks of trigger performance. For example, electrons and muons from inclusive W's and Z's, photons and jets at high p_t which overlap with measurements at lower $\sqrt{s}$, and missing E_t from Higgs $\rightarrow l^+l^-\nu\bar{\nu}$ or from SUSY decays can be used as benchmarks for the physics quanta of interest.

5.2.1 SDC Trigger Strategy

In the discussion which follows, the trigger strategy adopted by the SDC[3] is presented as an example of a strategy for high energy hadron collider experiments. SDC has adopted a three-level trigger architecture. For SDC, the strategy of Level 1 trigger is to identify "physics objects" and combinations of physics objects. The physics objects are electrons, muons, photons, jets, and "neutrinos," where a "neutrino" is signaled by missing E_t. Combinations of objects are dileptons, lepton plus missing E_t, *etc.* The Level 2 trigger then refines the identification of the physics objects, for instance by sharpening p_t and E_t cuts and by rejecting photon conversions. Finally, the Level 3 trigger identifies more complete physics signatures of physics processes, using the full event data and the capability of full analysis.

The SDC trigger strategy outlined here and below is still in a relatively early stage of development and can be expected to evolve significantly prior to implementation.

Table 3. Some Sample Physics Signatures of Interest at the SSC (from Ref. 3).

Physics Process	Mass Region (GeV)	Physics Signature
Associated Higgs Production	80 – 150	$W + H$, $t\bar{t} + H \rightarrow l\gamma\gamma$
Direct Higgs Production	130 – 180	$H \rightarrow ZZ^* \rightarrow 4l$
	180 – 800	$H \rightarrow ZZ \rightarrow 4l$
	500 – 800	$H \rightarrow ZZ \rightarrow 2l2\nu$
High Mass Boson Pairs		$Z\gamma \rightarrow l^+l^-\gamma$
Requires integrated luminosity of at least 50 fb^{-1} for complete studies	1–2 TeV	$W^+Z \rightarrow l^+l^+l^-\nu$
		$W^+W^+ \rightarrow l^+l^+$
Discovery of t Quark	$\lesssim$ 1 TeV	$t\bar{t} \rightarrow W^+W^- + X \rightarrow e^\pm\mu^\mp + X$
Mass Measurement of t Quark		
Sequential Dilepton Mode	$\lesssim$ 500	$t\bar{t}$, one $t \rightarrow Wb$; $W \rightarrow e\nu$; $b \rightarrow \mu + X$ the other $t \rightarrow 3$ Jets
Lepton + Jets + b-tag Mode	$\lesssim$ 500	$t\bar{t}$, one $t \rightarrow W + X$; $W \rightarrow l\nu$ the other $t \rightarrow Wb \rightarrow b + 2$ Jets
Non-standard t Decays		
Violation of τ Universality	$M_H \lesssim M_{top} - 15$	$t \rightarrow H^\pm b$; $H^\pm \rightarrow \tau^\pm\nu$; $\tau^\pm \rightarrow \pi^\pm + X$
Peak in 2-Jet Mass Distribution	$M_H \lesssim M_{top} - 25$	$t \rightarrow H^\pm b$; $H^\pm \rightarrow c\bar{s}$
Gluino and Squark Searches		
Missing-E_t + Jets	300 – 1000	$\tilde{g}\tilde{g} \rightarrow \not{E}_t^{miss}$ + 3–6 Jets
Like-Sign Dileptons	200 – 2000	$\tilde{g}\tilde{g} \rightarrow l^\pm l^\pm$ + 4 Jets
New Z Searches		
Discovery	$\lesssim$ 4 TeV	$Z' \rightarrow l^+l^-$
Width and Asymmetry	$\lesssim$ 2 TeV	$Z' \rightarrow l^+l^-$
Compositeness		
	$\Lambda \gtrsim$ 25 TeV	Inclusive Single Jet Spectrum

5.2.1.1 Electron Triggers

At Level 1, the first step in identifying electrons is to identify an electromagnetic shower, using calorimeter trigger towers with electromagnetic energy deposit above a threshold and with a ratio of hadronic energy deposit to electromagnetic energy deposit (E_{had}/E_{em}) below a small threshold. The E_{had}/E_{em} cut selects electrons by their longitudinal shower development; however, it also applies an implicit isolation cut, rejecting hadrons in the same tower and hadrons in nearby towers with cascades spreading into the trigger tower. Phototube discharges can be eliminated by demanding a coincidence of a hit trigger tower with a hit strip within the shower maximum detector within the trigger tower. Finally, electrons are distinguished from photons and π^o's by demanding the presence of a charged track associated with the shower, where a track is defined as stiff track segments in the outer tracker which point in ϕ to the trigger tower. In addition, there is the option to demand that the electron is isolated, by requiring that the sum of energy in surrounding trigger towers is below a threshold value.

At Level 2, three backgrounds to electrons must be rejected. Photon conversions, the dominant background at Level 2, can be rejected by demanding hits in the inner silicon layers along a track associated with the shower. Overlap of charged and neutral pions can be rejected by demanding a spatial match in ϕ of the associated track with the location of the shower as measured by the shower maximum detector. Finally, early showering charged hadrons, which are mostly rejected by the E_{had}/E_{em} cut at Level 1, can be rejected by a loose E/p cut. As in Level 1, the option to demand isolation also exists.

At Level 3, several additional electron identification criteria can be applied. E_t measurements can be sharpened, using the full calorimeter segmentation and resolution. Calorimeter and shower maximum energy profiles can be used. The spatial match between tracks and showers can be refined, using the shower maximum energy profile and finer tracker resolution. The measured calorimeter energy can be corrected for energy losses in cracks and inert materials. Finally, the rejection of photon conversions can be refined by performing additional track reconstruction.

For energetic electrons, not all of the above identification criteria are needed. In practice, trigger electrons will be identified by several parallel sets of criteria, with the strictest sets applied to the lowest energies, and with the loosest and most efficient set applied to the highest energy electrons.

5.2.1.2 Electron Trigger Rates

The example of an inclusive electron trigger, illustrates the general nature of event selection criteria which might be used. Figure 18 shows the results of an early study of electron triggers performed by Y. Sakai of KEK using a simple calorimetric model with fast shower simulation of QCD events generated by ISAJET. First the energy deposit in the electromagnetic section of a calorimeter tower of size approximately $\Delta\phi \times \Delta\eta = 0.2 \times 0.15$ is required to be above some threshold, probably in the range 20 to 40 GeV. The energy in the hadronic section is also required to be less than some fraction (20%) of the energy in the electromagnetic section. The resulting rate is shown by the solid curve in Fig. 18 as a function of E_t threshold. For example, rejection greater than 10^4 (*i.e.*, rate less than 10^4) is achieved for thresholds above 20 GeV. A stiff track segment with $P_t > 5$ GeV pointing towards the trigger cell in ϕ (i.e., with no z requirement) is then required. This criterion reduces the rate by about another factor of approximately 10, as shown by the dotted curve in Fig. 18. Finally, the trigger cell is also required to be isolated. That is, the energy in nearest neighbor cells, electromagnetic and hadronic, is required to be less than 20% of the energy in the trigger cell. Rejection greater than 10^6 will then be achieved for all energies greater than about 12 GeV, as shown by the dashed curve in Fig. 18. As explained in the section above, further rejection can be achieved via other selection criteria.

5.2.1.3 Photon Triggers

At Level 1, the photon trigger is very similar to the electron trigger, with the exception of demanding a charged track. Electromagnetic showers are identified by calorimeter trigger towers above threshold, with a small ratio of hadronic energy deposit to electromagnetic energy deposit. Photomultiplier discharges are rejected by demanding a hit in the shower maximum detector within the trigger tower. There is also the option to demand that the photon is isolated.

At Level 2, many π^o's can be rejected by examining the shower profile in the shower maximum detector. As in Level 1, the option to demand isolation exists.

At Level 3, E_t measurements can be improved by using the full calorimeter segmentation and resolution and by correcting for cracks and inert materials in the calorimeter. Finally, advanced pattern recognition can be performed on shower profiles in the calorimeter and shower maximum detector.

Of course, electron triggers are distinguished from photon triggers by the presence of a charged track. Consequently, photon triggers will not be as effective at background rejection as electron triggers, and will therefore have higher E_t thresholds.

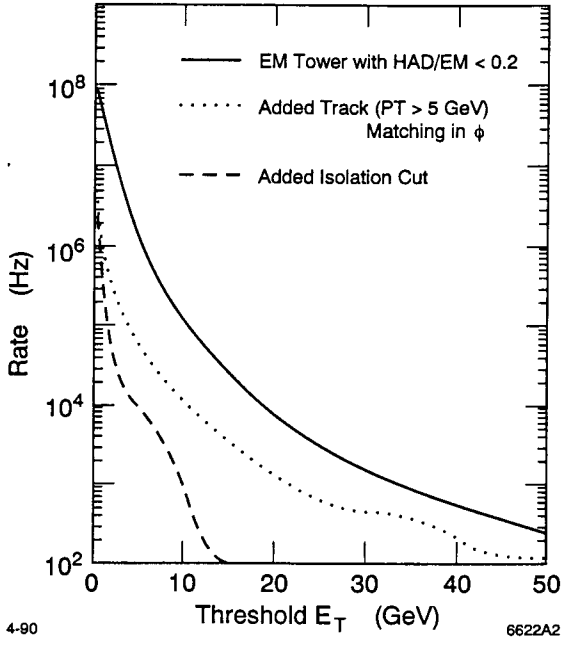

Fig. 18: Plot of trigger rate vs. transverse energy threshold for various electron selection criteria.

5.2.1.4 Muon Triggers

At Level 1, muon candidates are identified by the presence of a stiff track segment in the outer chambers of the muon system. Hadronic backgrounds are greatly reduced by the absorber present in the hadron calorimeter and in the muon toroids. Tracks from the remnants of hadron cascades in general do not point back to the origin, as do stiff muon tracks. The p_t of the muon candidates are determined by the angle (in the θ coordinate) of the track in the muon chambers after the track is bent by the toroidal magnet. Low energy muons either range out in the absorber or are eliminated by a p_t cut. Because the muon chamber drift time is long, of order one microsecond, the bunch crossing from which the muon candidate originates is tagged by scintillators in the muon system. In addition, there is the option at Level 1 to demand a stiff track segment in the outer tracker as a means of sharpening the p_t cut or of rejecting accidental stiff tracks in the muon system. Finally, there is the option to apply an isolation cut by demanding that the energy in calorimeter towers surrounding the interpolated muon track is below threshold.

At Level 2, the most effective criterion in selecting high p_t muons is to sharpen the muon momentum threshold by linking the muon segment found at Level 1 to a central tracker segment. The central tracker provides improved momentum resolution, rejecting muons at lower p_t from appearing to be above threshold because of resolution effects. As in Level 1, the option to demand isolation also exists.

At Level 3, muon identification can be refined by performing full three-dimen-sional track reconstruction.

5.2.1.5 Jet Triggers

At Level 1, jets are identified and measured by localized calorimeter energy above threshold. Two techniques are possible. The first identifies "seed" towers by demanding that a relatively small "supertower" of trigger towers, say $\Delta\phi \times \Delta\eta = 0.2 \times 0.2$ or 0.4×0.4, have an energy sum above threshold, thus triggering on the core of the jet. The second technique triggers on the full energy of the jet by containing it in a large supertower, *e.g.*, $\Delta\phi \times \Delta\eta = 1.6 \times 1.6$, where the supertowers are arranged in overlapping grids such that a jet cannot fall on the boundaries of all grids. Studies suggest that event pileup in the large supertowers is not a significant problem, even for luminosities as high as 10^{34}, well above design luminosity. For fixed trigger rate, large supertowers achieve lower thresholds with better efficiency than do small supertowers.

At Level 2, jet E_t measurement can be improved by using clustering or fixed-cone algorithms. The jet algorithm at Level 2 should be of a similar nature to the one to be used later offline in order to simplify the analysis of efficiencies.

At Level 3, E_t measurements can be further refined using the full calorimeter segmentation and resolution, correcting for cracks and inert material in the calorimeter, and using refined jet clustering or cone algorithms.

5.2.1.6 "Neutrino" Triggers

At Level 1, neutrinos and neutrino-like particles are identified by determining the missing E_t of the event. The missing E_t is computed using only calorimeter trigger towers above some low threshold. Doing so minimizes the effects of detector noise and of soft pileup.

At Level 2, the missing E_t measurement can be refined by correcting for muons. In addition, it is possible to demand that the direction of the missing E_t vector does not point towards a dead region in the detector. An additional option is to recompute the missing E_t using only energy that has been reconstructed in jets. Studies suggest that this method has the best missing E_t resolution.

At Level 3, the missing E_t measurement can be further refined by using the full calorimeter segmentation and resolution and by correcting for cracks and inert materials in the calorimeter.

5.2.2 Role of the Calorimeter in SSC Triggers

As can be seen from the above descriptions of trigger selection criteria, the calorimeter plays a central role in most SSC triggers. The calorimeter provides the most effective way to accomplish fast, efficient reduction of event rate by selecting interesting high-p_t events and identifying electrons, photons, jets, and missing E_t. It can even be used to assist in muon triggers. The data inputs and algorithms of the SDC calorimeter trigger are sketched here as an example of defining inputs and algorithms during the trigger design process.

The data needed from the calorimeter by the trigger in SDC is defined as coming from trigger towers of size $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$. The electromagnetic and hadronic compartments of the trigger tower will be separately summed. The sums will be digitized every beam crossing with a twelve-bit dynamic range, although an eight-bit non-linear response, or perhaps even an eight-bit linear response, is adequate. The two eight-bit sums from a trigger tower will be output each crossing on a 1 Gbps fiber which will take the trigger data to the trigger processors. SDC will use this same calorimeter trigger data for both its Level 1 and Level 2 triggers. The 1 Gbps fiber optics employed is standard for all Level 1 trigger data links.

The data needed from the shower maximum detector at Level 1 is simply one bit per crossing per $\Delta\phi \times \Delta\eta = 0.1 \times 0.2$ region indicating that one or more of the shower maximum strips in that region is above threshold. At Level 2, data from individual shower maximum strips will be provided to the trigger, although it has not yet been determined whether improved resolution warrants providing the trigger with pulse height information, rather than simply lists of hit strips.

The SDC calorimeter trigger processor at Level 1 searches trigger towers for electromagnetic showers, sums E_t in overlapping $\Delta\phi \times \Delta\eta = 1.6 \times 1.6$ regions, sums total E_t, calculates missing E_t, and checks isolation by searching energy cones. At Level 2, the processor uses the same data as at Level 1 to recompute E_t for electrons, jets, missing E_t, and isolation, using clustering or fixed-cone algorithms. It also provides information for a loose E/p cut.

The SDC shower maximum detector trigger processor at Level 1 associates a hit region of the shower maximum detector with a hit trigger tower. At Level 2, the processor associates the ϕ-position as measured by the shower maximum detector with the projected position of a track in the central tracker. At Level 2 the processor may also perform loose shower profile cuts for electron and photon candidates.

5.2.3 Data Paths to Trigger Processors

The Level 1 and Level 2 triggers of SDC are each organized as a set of local (or subsystem) triggers which process input trigger primitives and which provide outputs to global Level 1 and Level 2 trigger processors. The local triggers process the data from a single subsystem. The global triggers combine the trigger data from the local triggers.

Figure 19 outlines the organization of the trigger processors, highlighting the data required by the three-level trigger of SDC and the data paths to the trigger processors. In this figure, the front-end electronics of the detector subsystems are shown in the left column. The other columns show "Level 1 local trigger processors," "Level 2 local trigger processors," and "data acquisition buffers." These three columns of local processors and local buffers connect vertically to the "global Level 1 and Level 2 triggers" and to the data acquisition system and Level 3 trigger. The contents of the local processor boxes in the figure identify the resolution of the outputs of the local trigger processors to the global triggers. For instance, the local Level 1 trigger processor for the central tracker outputs track segments in 64 ϕ bins and 2 η bins with each segment represented by two bits of data.

The data paths from the front-end electronics to Level 1 local trigger processors and to Level 2 trigger local processors are shown in the figure as horizontal connections. For instance, the electromagnetic and hadronic calorimeter front-end electronics transmit trigger primitives to the local Level 1 calorimeter trigger. The local Level 2 calorimeter trigger obtains its input data from the local Level 1 calorimeter trigger. All Level 1 local processors receive their input data from the front-end electronics. Level 2 local processors receive their input

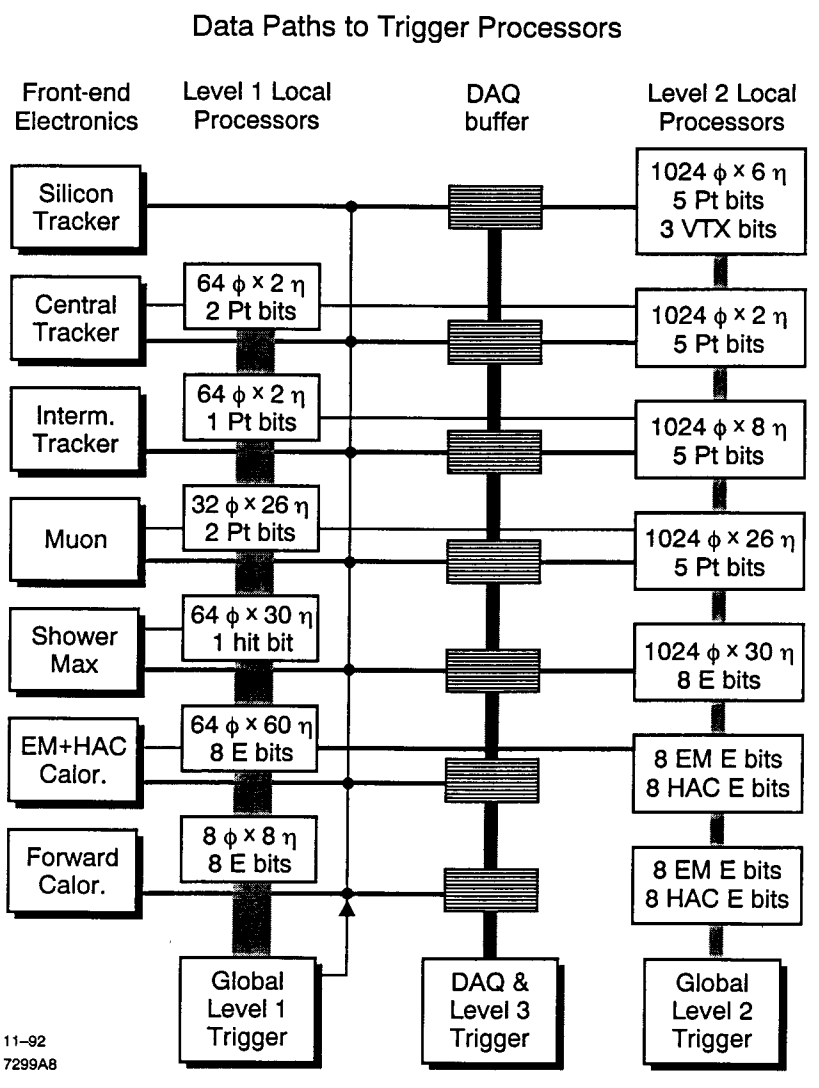

Fig. 19: Architecture of trigger processors for the SDC Experiment, showing trigger data paths among front-end electronics and processors and defining granularity and resolution of the data used by each processor.

data from either the Level 1 local processors or from the front-end electronics, or from both, according to whether the data from Level 1 is adequate for the Level 2 algorithms. The definition of the data requirements of each level is based on study of the physics processes of interest and of the background processes.

5.2.4 Examples of Level 1 Trigger Thresholds in SDC

Based on extensive simulation of detector response and potential Level 1 trigger algorithms, Table 4 summarizes the Level 1 trigger thresholds expected in SDC for some inclusive triggers. Study of Levels 2 and 3 are continuing.

Table 4. Examples of SDC Level 1 Trigger Rates.

Trigger	Threshold
Electron	20 GeV
Photon	30 GeV
Muon	20 GeV
Jet (1.6x1.6 sum)	140 GeV
Missing E_t	80 GeV
Dielectrons	10 GeV
Diphotons	20 GeV

5.2.5 B Physics Triggers in a Hadron Collider

The requirements for a hadron collider experiment focusing on B physics are quite different from those described above for high-p_t physics. Whereas, the calorimeter trigger forms the basis of triggers for high-p_t physics, B physics with large event samples places a premium on track finding in the early levels of the trigger. For instance, B physics proposals for the Fermilab collider and for the SSC have discussed a very challenging topology trigger which identifies the relatively stiff tracks which obtain their momentum from the mass and p_t of parent B mesons. For instance, this trigger might require at least one track with $p_t > 3$ GeV or at least two tracks with $p_t > 2$ GeV. The logical OR of these criteria might result in an enhancement of a factor of 50 in the B sample. An alternative trigger is to tag B decays with electrons, where the p_t threshold discussed for electron candidates is as low as 2 GeV. Clearly, these are very challenging triggers in the difficult environment of a hadron collider.

5.3 TRIGGERS AT PEP II

Unlike triggers for hadron colliders, the trigger at an asymmetric B-factory, such as PEP II, must retain essentially all physics events, *i.e.*, all annihilation processes, as well as a fraction of two-photon physics. The technical challenge is that the trigger must reduce the rate by about six orders of magnitude, from 2.5×10^8 crossings per second to about 100 triggers per second. Although the physics challenges are quite different, the technical challenge of large rate reduction at PEP II is sufficiently similar to the SSC that the architectural and design approaches can also be similar.

The annihilation cross sections and rates are shown in Table 5 for the upsilon resonances. The physics rates are between 16 and 72 Hz at luminosity of 3×10^{33}. The dominant physics signatures at the $\Upsilon(4s)$ are listed below :

- $e^+e^- \rightarrow B\overline{B}$,
- $e^+e^- \rightarrow e^+e^-$ (Bhabha) ,
- $e^+e^- \rightarrow \mu\mu$,
- $e^+e^- \rightarrow \tau\tau$,
- $e^+e^- \rightarrow e^+e^-X$ (two-photon) .

The most difficult class of events upon which to trigger are $\tau^+\tau^-$ pair events in which each τ decays into a single charged prong with no neutrals. This process demands a trigger on two charged tracks which are not colinear, not necessarily electromagnetic showering, and not necessarily penetrating.

Table 5. Cross Sections and Production Rates of the Υ Resonances.[6]

Resonance	Mass(GeV)	R_{had}	R_{tot}	σ_{pt}	σ_{tot}	Rate at 10^{34}	at 3×10^{33}
1s	9.46	22	25	0.97 nb	24 nb	240 Hz	72 Hz
2s	10.02	10	13	0.87 nb	11 nb	110 Hz	33 Hz
3s	10.36	7	10	0.81 nb	8 nb	80 Hz	24 Hz
4s	10.58	4	7	0.78 nb	5.5 nb	55 Hz	16 Hz

5.3.1 PEP II Trigger Rates

Table 6 shows a range of acceptable trigger times and rates for each level in a three-level architecture. The values shown in Table 6 are loosely based upon experience with trigger processing techniques in prior experiments and reflect the flexibility which is available in designing the trigger. The approximately one microsecond decision time for Level 1 corresponds to the minimum time to collect trigger signals, perform rudimentary processing, and distribute strobes. More complex processing can be done by hardware in ten or more microseconds by Level 2. Ten microseconds processing time limits the input rate to 100 kHz from Level 1. Lower rates from Level 1 would afford Level 2 additional processing time. The lowest conceivable rate from Level 1 (as discussed below) is 1 kHz, which would afford Level 2 as much as one millisecond processing time. Such a low rate might eliminate the need for the data acquisition system to buffer multiple events during the Level 2 decision, or it could possibly eliminate the need for a hardware Level 2 trigger preceding the Level 3 trigger. Order 100 Hz, which is near the physics rate, is the lowest conceivable rate into Level 3. Since the Level 3 trigger is usually a collection of general-purpose microprocessors, its input rate is bounded at approximately 1 kHz, which would allow each processor only one millisecond times the number of processors to select events.

Table 6. PEP II Trigger Levels and Rates.[6]

Parameter	Level 1	Level 2	Level 3
Avg. Time between Decisions	4-12 nsec	10-1000 μsec	1-10 msec
Average Decision Time	$\sim$ 1 μsec	$\sim$ 100 μsec	10^2-10^3 msec
Input Rate (Hz)	2.5×10^8	10^3-10^5	10^2-10^3
Expected Rejection	10^4-10^5	10-10^2	1-10
Output Rate (Hz)	10^3-10^5	10^2-10^3	$\sim$ 100

5.3.2 PEP II Trigger Inputs

Figure 20 illustrates detector information which is available to and which may be utilized by each trigger level of an experiment at PEP II. At Level 1, calorimeter triggers will operate on energy deposit in trigger towers and on total energy. Tracking triggers at Level 1 will operate on either hit cells or track segments in the central drift chamber. At Level 2, more detailed drift chamber data, such as drift times, will be used. Vertex information based on ORs of

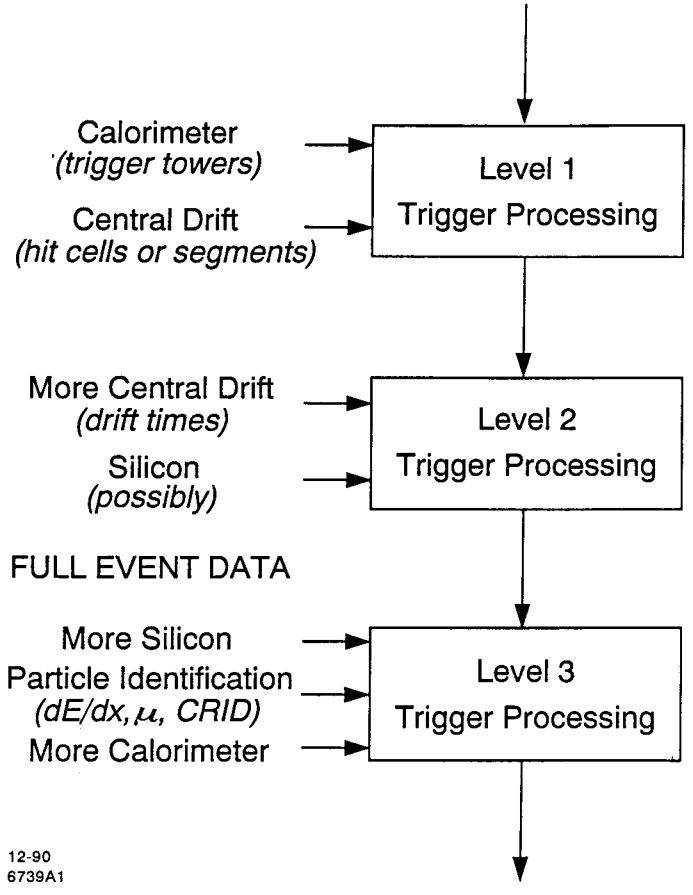

Fig. 20: Illustration of the detector information which is available to and which may be utilized by each trigger level of an experiment at PEP II (from Ref. 6).

adjacent strips or pixels may also be used. Data from the finely segmented silicon system may be difficult to access on the time scales of the Level 2 trigger, and are impossible to access in a Level 1 trigger. At Level 3, data from the full digitized event is available. This data includes vertex information from the silicon system, particle identification information such as dE/dx, muon system hits, ring imaging Cherenkov data, and complete calorimetry data. Time-of-flight data, if it exists, would be available as early as Level 1. The nature of data used by each level and how the data is used will be further described in sections which follow.

5.3.3 PEP II Trigger Conditions and Efficiency

The prompt triggers at PEP II are expected to be elemental, based simply on numbers of charged tracks or on calorimetric energy deposit. The most challenging final states upon which to trigger are those with both low charged multiplicity and low visible energy, such as the $\tau^+\tau^-$ and two-photon processes.

An adequate charged particle trigger would be a logical OR of a two-track trigger (referred to as 2T in Table 7) and a one-and-a-half-track trigger (referred to as 1.5T). The two-track trigger requires at least two tracks which penetrate to the outermost drift chamber layer. The one-and-a-half-track trigger requires at least one track which penetrates to the outermost drift chamber layer and at least one track which may pass through only the inner drift chamber layers. The one-and-a-half-track trigger is designed to improve acceptance for two track events, which may be boosted into unfavorable topologies. Single track triggers are expected to result in unacceptably high rates due to backgrounds.

The calorimetric trigger could be a logical OR of a total energy trigger (referred to as E in Table 7) and a two-supertower trigger (referred to as 2ST). The total energy trigger would be based on the visible energy in the calorimeter (barrel and endcap), which would be required to be above a threshold, typically of order 1 to 2 GeV. A low channel-by-channel threshold would be applied to each calorimeter tower prior to summing. The two-supertower trigger would require at least two contiguous groups of towers (which define a supertower) to have a combined energy deposit greater than one-half of minimum ionizing, or about 100 MeV. In order to reduce background due to cosmic rays and to beam gas, and other beam related noise, it may be necessary to implement a topological requirement in the calorimetric trigger, for instance to require a certain amount of azimuthal symmetry to the energy deposit.

The efficiency of these triggers has been studied with the ASLUND Monte Carlo by the SLAC B-factory detector working group.[6] Our study included both hadronic final states due to $e^+e^- \rightarrow B\overline{B}$ for which efficiencies are expected to be high and the presumably challenging case of $\tau^+\tau^-$ pairs. For this study, calorimeter response was simply modeled by assuming that for photons

and electrons all energy is visible, for muons exactly the energy deposit of one minimum-ionizing particle (mip), *i.e.*, 200 MeV, is visible, and for hadrons the visible energy is evenly distributed between 1 mip and the total energy, or zero and the total energy depending on whether the hadron is charged or neutral respectively. Supertowers were defined with angular size $\Delta\phi \times \Delta\theta = 2\pi/24 \times \pi/12$ and were called "hit" if the visible energy exceeded 100 MeV. For details of the topological trigger (referred to as SHC), see Ref. 6. A plot of the efficiency of the total energy trigger for $\tau^+\tau^-$ pairs and for the "CP" final state $B^0 \to \psi K_S$ is shown in Fig. 21 as a function of the total energy threshold. Table 7 summarizes our results for several triggers. The efficiency is seen to be high. Figure 21 shows that the efficiency for $B^0 \to \psi K_S$ events is near 100% for reasonable total energy thresholds; however, the efficiency for $\tau^+\tau^-$ pairs falls rapidly as threshold increases, starting below 1 GeV threshold where the efficiency is 88%. Nonetheless, as Table 7 shows, the two supertower trigger 2ST is 95% efficient for $\tau^+\tau^-$ pairs. Recall that the first four triggers in Table 7 would be ORed, except for the final topological trigger which could be required for any calorimeter trigger.

Table 7. Selected Trigger Efficiencies at PEP II.[6]

Trigger	$\tau^+\tau^-$	$B^0 \to \psi K_S$
2T	0.530	0.997
1.5T	0.730	1.000
2ST	0.950	1.000
E	0.883	1.000
SHC	0.843	1.000

5.3.4 Expected Backgrounds at PEP II

Four sources of background to the trigger will be present at an asymmetric B-factory such as PEP II: cosmic rays, beam gas scatters within the detector volume, synchrotron radiation, and lost beam-particles. The flux of cosmic rays striking any part of the calorimeter or inner detector will be approximately 3 kHz, 90% of which will be single muons near minimum-ionizing and which may be eliminated by loose geometrical cuts. The beam-gas background rate depends on the storage ring vacuum and may be between 100 Hz and 10 kHz.[6] Preliminary calculations[6] of synchrotron radiation and lost-particle backgrounds indicate that lost-particle backgrounds are dominant, even for low energy photons.

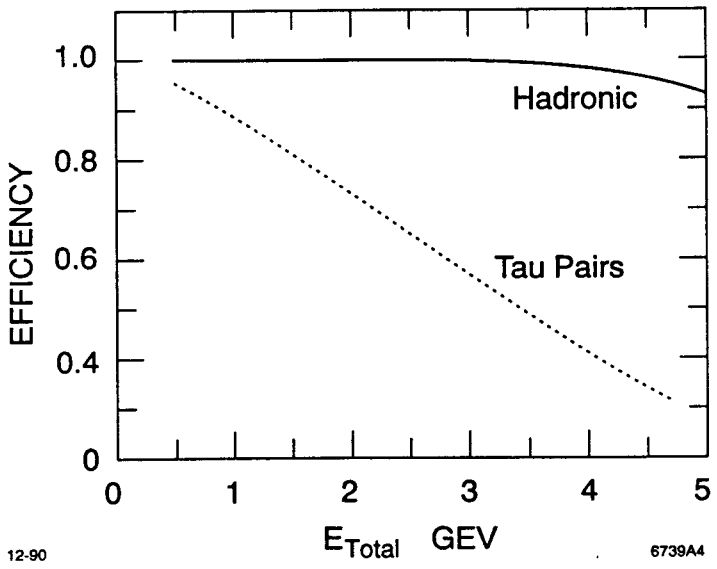

Fig. 21: Plot of the efficiency of a total energy trigger at PEP II for $\tau^+ \tau^-$ pairs and for the "CP" final state $B^0 \to \psi K_S$ as a function of the total energy threshold (from Ref. 6).

Results of a preliminary study[6] of lost-particle backgrounds are summarized in Table 8 for luminosity of 3×10^{33}, although the calculated background rates are extremely dependent upon the accelerator lattice. Contributions both from random overlap of two different background rays and from two significant tracks or energy deposits from a single ray were considered; however, "fake" charged tracks due to overlap of background hits in the drift chamber were not considered. Note that charged particles with $p_t < 100$ MeV (assuming one meter calorimeter radius and 7 KGauss magnetic field) will not reach the calorimeter; so the trigger level is likely to be at or somewhat above that value. A minimum-ionizing track will deposit about 200 MeV in a cesium iodide calorimeter or about 350 MeV in a liquid krypton calorimeter.

Table 8. Charged Particle and Calorimeter Trigger Rates Due to Lost Beam Particles.[6]

Trigger	Cut (MeV)	Random Overlap Rate (kHz)	Rate from one ray (kHz)	Total Trigger Rate (kHz)
$\geq$ 1 Supertowers	50	–	290	290
	100	–	52	52
	200	–	10	10
$\geq$ 2 Supertowers	50	85	31	116
	100	3	2	5
	200	0.1	< 5	< 5
$\geq$ 1 Track	60	–	20	20
	100	–	4	4
$\geq$ 2 Track	60	0.4	0.3	0.7
	100	0.02	< 0.8	< 0.8

The background rates from these four sources are manageable; however, they are sufficiently close to the Level 1 target of 10 kHz to be of concern. In particular, the backgrounds from lost beam particles must be carefully evaluated as the machine design evolves.

5.3.5 Level 1 Triggers at PEP II

In order to reduce rates to the 10 kHz level, the Level 1 trigger at PEP II must eliminate backgrounds due to low energy (below p_t of 50-100 MeV) tracks and photons from synchrotron radiation and from showers of lost beam particles. In addition, it must eliminate most of the background from lost beam particles by requiring that there be two particles (charged or neutral) above the p_t cut. The output of the Level 1 trigger will then be dominated by cosmic rays and by coincidences of secondaries from showers of one or two lost beam particles. The actual p_t cut implemented will depend upon the resolution of the trigger and upon background rates, but should be sufficiently low to maintain efficiency for two prong final states of annihilation events.

5.3.5.1 PEP II Level 1 Calorimeter Trigger

Physics requires that the electromagnetic calorimeter of the detector at PEP II be efficient for low energy photons from π^o decays. Consequently, the calorimeter will have very good signal-to-noise for individual minimum-ionizing particles, approximately 400:1 signal-to-electronics noise for a single calorimeter cell of either cesium iodide or liquid krypton. This excellent performance would enable a very efficient trigger for events with two or more minimum ionizing particles, without the use of charged particle tracking. The dominant rate for such a trigger could be cosmic rays which pass through any portion of the calorimeter, a background rate of approximately 3 kHz. As mentioned above, this rate could possibly be reduced by requirements on the topology of energy deposit in the calorimeter, and it can be substantially reduced by requiring tracking in some subsequent trigger level.

A Level 1 calorimeter trigger would likely consist of three parallel components, combined in a logical OR: the two supertower trigger (2ST), the total energy deposit trigger (E), and a trigger on the topology of supertower energy (SHC). Here we describe a two supertower trigger. For more information on the other calorimetric triggers, see Ref. 6.

The two supertower trigger is accomplished by summing nearby calorimeter towers into "supertowers" and then counting the number of supertowers above threshold. Groups of contiguous calorimeter towers are summed into supertowers of about 5×5 individual towers. A supertower is illustrated in Fig. 22. Summing reduces the number of calorimeter trigger channels to about 400 from the order of 10^4 individual towers. It also reduces the probability that the energy of a particle will be split between two trigger channels. The signal-to-noise for a minimum-ionizing particle in a supertower will still be very good, as high as 80:1, enabling an efficient threshold of order one-half the minimum-ionizing signal. To further improve efficiency at boundaries of supertowers, where energy

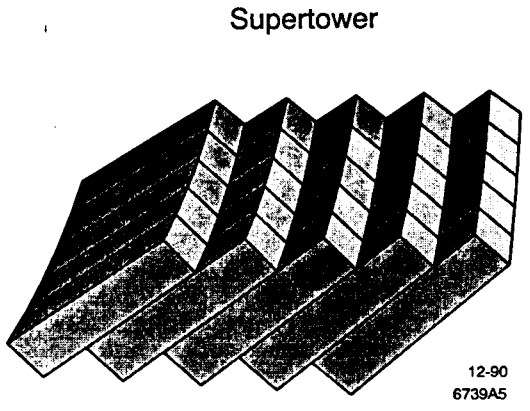

Fig. 22: A calorimeter trigger supertower at PEP II formed from a 5×5 array of calorimeter cells.

may be split between two supertowers, overlapping sets of supertowers may be formed. We refer to a complete set of supertowers covering the calorimeters as a "layer." The two supertower trigger requires at least two non-adjacent supertowers above threshold in a single "layer" as a means of providing a trigger on two particles without double-counting a single track. It is unlikely that pileup of low-energy photons will accumulate enough energy in a single supertower (or pair of supertowers) to satisfy a threshold at half the energy deposit of a minimum-ionizing particle. The single tower rate due to electronic noise alone is negligible. The rate of two supertowers in coincidence is even less. For supertowers above threshold, the time of the event can be determined using techniques described in Section 5.1.1. The timing resolution of a single minimum-ionizing particle in a supertower is expected to be a few nanoseconds, and can be improved with signal shaping optimized for timing, if necessary. Two supertowers providing a trigger candidate must identify the same crossing. Identification of the bunch crossing also contributes to knowledge of the drift chamber t_0 for use by the Level 2 tracking trigger.

5.3.5.2 PEP II Level 1 Tracking Trigger

The Level 1 tracking trigger must identify charged particle tracks above 50–100 MeV p_t which originate from the region in which the beams interact. It discriminates against low energy tracks which arise from showers of lost beam particles, and from combinations of random hits from synchrotron radiation, other low energy photons, or out of time hits from other bunch crossings. The volume along the beam line from which tracks appear to originate is referred to as the "fiducial" volume for the trigger. The size of the fiducial volume determines the rate of accepted cosmic rays and tracks from lost beam particles. The radius of the fiducial volume is determined by the resolution of the track finder and by the radius of the inner-most tracking layer used by the trigger. Along the beam line, *i.e.*, in z, the fiducial volume at Level 1 is determined by the active length of the inner-most tracking layer. Consequently, the fiducial volume at Level 1 will be a few centimeters in radius and a couple meters in length. Less than 100 Hz of cosmic rays, but a few kHz of high-p_t tracks produced by lost beam particles, will traverse this volume.

The Level 1 tracking trigger consists of two steps: recognizing segments of tracks in the $r-\phi$ projection of superlayers of the chamber, and linking segments into projected tracks. The segment finding step will eliminate most out-of-time and other random hits and will provide some p_t threshold. The segment linking step will further suppress accidental combinations of random hits and will provide a well-defined p_t threshold. The algorithms by which these steps, particularly segment finding, will depend on whether the chamber superlayers are composed of small cells or of jet cells. Figure 23 shows schematically for both

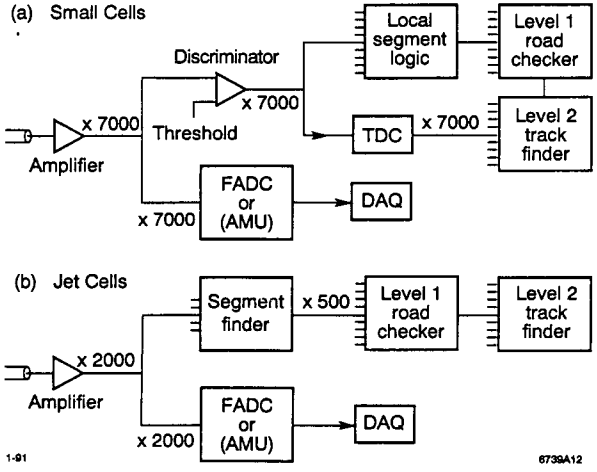

Fig. 23: Simplified block diagram of logic at PEP II for tracking triggers for drift chambers with (a) small cells and (b) large cells (from Ref. 6).

small cells and jet cells the logical steps in the tracking trigger. For either cell type, discriminated chamber signals could be formed into track segments locally and then linked into tracks by a "Level 1 road checker." In both cases, the output of the Level 1 road checker could be used in the "Level 2 track finder." In the case of small cells, the Level 2 track finder is likely to also need to use drift times from the TDCs. In the case of jet cells, the additional position resolution from drift time is already available to the "segment finder." For small cells, using the half-cell offset of adjacent layers of chamber superlayers, as explained in Section 5.1.1, allows event time determination. For both cell geometries, Fig. 23 also shows a flash ADC or analog memory (FADC or AMU) which is read out for particle identification but which is not used in the trigger. Reference 6 describes possible implementations of triggers for each scheme. Since the writing of Ref. 6, the authors' concept of the trigger for the jet cells has advanced considerably.

The Level 1 tracking trigger may need to demand either at least one track or at least two charged tracks, depending on its resolution and upon beam-related background rates. If the Level 1 calorimeter trigger is efficient for events with two minimum ionizing particles and meets rate requirements, then the Level 1 tracking trigger need only be sufficiently efficient to provide a measure of the efficiency of the calorimeter trigger.

5.3.6 Level 2 Triggers at PEP II

The trigger rate into Level 2 at PEP II will be dominated by cosmic rays and by coincidences of secondaries from showers of one or two lost beam particles. These backgrounds have similar two-prong topologies as some interesting physics processes, such as production of $\tau^+\tau^-$-pairs; however, background tracks in general will not originate at the beam interaction point. Consequently, the principal manner by which the Level 2 trigger can reduce the trigger rate is by restricting the fiducial volume for tracks in $r-\phi$ and in z. This is accomplished using tracking information of higher resolution than was used in Level 1. With the exception of triggers on totally neutral events, calorimetric information is not important at Level 2 since it was already used at Level 1. With adequate resolution near the vertex in Level 2, the trigger rate from Level 2 may be dominated by beam-gas backgrounds within the vertex volume.

The rate of cosmic rays into Level 2, which may be as high as 3 kHz, is based on the large fiducial volume accepted by the Level 1 calorimeter trigger. This rate may be easily reduced by requiring more than one charged particle track within a smaller fiducial volume. For instance, a fiducial volume comparable to that used in the Level 1 tracking trigger, which is a few centimeters in radius and from ten to a few tens of centimeters in length, would result in a cosmic rate of order 50 Hz, or roughly the same as the physics rate with 10^{34} luminosity. Further reduction in rate is possible either with further reduction in the Level 2

fiducial volume or by use of time-of-flight information; however, further reduction at Level 2 is not critical. The cosmic rate is within the overall trigger rate goal of 100 to 1000 Hz for Level 2 and is less than other rates at that level. If further reduction in the cosmic rate is necessary, reduction at Level 3, where more data and resolution is available, may be more suitable.

The background rate into Level 2 from charged tracks produced by lost beam particles is expected to be between hundreds of Hertz and about 4 kHz (see Table 8). It also can be sharply reduced by restricting the fiducial volume, particularly in z. The rate is expected to drop roughly in proportion to reduction in z of the fiducial volume. Thus, more than one order of magnitude reduction is possible by reducing the z acceptance to order five centimeters. Further reduction in the z acceptance will be limited by the length of the beam bunches in z. The fiducial volume cannot be made smaller than the luminous region. Consequently, to further reduce backgrounds from lost beam particles it will be necessary to reduce the fiducial volume in $r - \phi$. Such reduction may be possible using either drift time measurements or by requiring that tracks have associated hits in the silicon vertex detector. The success with which these backgrounds can be reduced by reduction of acceptance in $r - \phi$ has not yet been studied.

Background from inelastic scattering of beam particles with residual gas within the vacuum at the interaction point (beam-gas) will contribute a small fraction of the trigger rate into Level 2, perhaps in the 50 to 100 Hz range if the PEP II vacuum is comparable to previous e^+e^- machines. However, if the vacuum at the PEP II interaction point is significantly worse, due to the higher beam currents illuminating the smaller beam pipe, residual beam-gas rates may be serious, of order hundreds of Hz. The background rate could be reduced by restricting the fiducial volume, but acceptance cuts in $r - \phi$ are ineffective since the background tracks originate along the beam line. Cuts on acceptance in z are limited to the length of the bunches. Consequently, reduction in beam-gas rate by Level 2 is limited to between one and two orders of magnitude, with rates of 1 to 10 Hz with a good vacuum. It may then become necessary to make additional topology cuts at Level 2. The asymmetric beam energies make cuts on momentum balance in z unusable; however, cuts on p_t balance, acoplanarity, and if necessary multiplicity may be effective. Any such cut will reduce efficiency for detection of $\tau^+\tau^-$ pairs, but will probably have little effect on B events. The rate of beam-gas background and techniques to reduce the rate have not yet received ample study.

5.3.7 Level 3 Triggers at PEP II

The trigger rate into Level 3 at PEP II will have approximately equal contributions of cosmic rays, backgrounds from lost beam particles, and physics events. In addition, there will be a contribution from beam-gas events which may be comparable to the other contributions, or which may be either larger or smaller depending on the vacuum at the interaction point. If all other backgrounds into Level 3 are comparable to the physics rate, then it is reasonable to record all events without further selection at Level 3. However, a Level 3 trigger can be implemented to further reduce background or to battle beam-gas backgrounds.

The Level 3 trigger has available the same data and the same processing engines which are available to offline analysis. It differs from offline analysis in that refined detector constants are not available and in that interactive choice of event selection criteria is not possible. However, given enough processing power in Level 3, it will be capable of performing online nearly any further event selection done offline. In particular, Level 3 will be capable of doing refined tracking and some vertexing to further reduce backgrounds which do not come from the interaction point. It will also be capable of applying complex topological cuts which may be necessary to eliminate beam-gas backgrounds which originate within the interaction volume. The residual backgrounds expected from Level 2 are all low multiplicity backgrounds which Level 3 should be able to reconstruct and analyze with modest computing power.

5.4 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 10^5 "usable" gates per chip, and will offer more in the future. Silicon compilation offers advanced cell libraries and development tools for semi-custom 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, and hence offers more patterns than possible with standard memory look-up techniques.

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 DSPs 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, particularly for two-dimensional detectors such as calorimeters, pad chambers, and pixel detectors. Some of our local pattern recognition problems, such as track segment finding and energy clustering, may be sufficiently simple to map onto neural nets of realizable scale. Alternatively, neural nets may serve as a paradigm for some application of massive parallel processing. Fine grained parallelism, as provided by the Associative String Processor (ASP) or by the Connection Machine, is also under investigation. These processors may provide a way of matching the granularity of the processing to the granularity of the detector. Additionally, data-driven pipelined processors which are evolutions of the one used in FNAL E690 could provide tremendous computational power to algorithms which can be defined well and which do not need the full flexibility of high-level language programmable processors. The computational power of these processors can be applied to pattern recognition problems as well as to performing calculations such as track fits.

Special-purpose processors, including 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 will fade as DSP and RISC cores are embedded in custom circuits and as custom coprocessors are attached to general-purpose CPUs. 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.

6. Summary and Prospects

Many common issues at future e^+e^- and hadron colliders give rise to many common requirements for trigger and data acquisition systems in the two environments. The common issues include the search for rare physics processes, high luminosities, and frequent beam collisions. In fact, the search for rare physics processes gives rise to the need for high luminosity, and the need for high luminosity demands frequent beam collisions. In addition, highly segmented detectors will contribute to high data rates.

The common requirements for front-end systems include highly integrated electronics, pipelined data buffering during trigger processing, and simultaneous analog and digital signal processing. For trigger systems, the common requirements include pipelined trigger processors, bunch-crossing identification, and highly sophisticated processors with substantial rejection. The common requirements for data acquisition are for high data bandwidths, extensive parallelism, and massive computing resources. In addition, thoughtful design of these large electronics systems is required.

The common requirements of the two environments give rise to the potential for similar architectural solutions for trigger and data acquisition. Nonetheless, differences in the physics goals and in the detector elements of each experiment give rise to important differences in the detailed requirements, which will lead to differences in the implementations. Indeed, other innovative architectural solutions are also possible for many applications.

Much development remains on the architecture and details of these future systems. The solutions implemented on experiments in the years ahead will certainly differ from those sketched in these lectures. In fact, designs of the SSC and PEP II examples have evolved significantly between the time that these lectures were delivered and the time that this document was written. Several years will elapse before experiments will be running at PEP II or SSC. Future experiments should use this time to build upon the rapid advances in the electronics, computer, and communications industries to explore new simple and cost-effective solutions. Much work remains in developing the architecture and the details of these future systems.

Acknowledgments

The concepts presented in these lectures are the product of a very large number of scientists studying electronics for future colliders over a number of years. The contributions of others are far too many to recognize individually; nonetheless, the author would like to acknowlege all contributions to the work presented in these lectures. Most notably, the examples shown result from the work of the SDC Electronics Group and the SLAC Asymmetric B-factory Trigger and Data Acquisition Group.

This work is supported in part by Texas National Research Laboratory Commission grants RGFY9225 and RGFY9266 and by Department of Energy grant DE-FG03-91#R40679. The author would particularly like to express his appreciation to the Texas National Research Laboratory Commission for support of the study reported in these lectures.

References

1. M. Breidenbach, "Data Acquisition for High Energy Physics Experiments," *Proceedings of the Fourteenth SLAC Summer Institute on Particle Physics*, SLAC Report 312, Stanford, CA, 1986.
2. R. K. Bock, H. Grote, D. Notz & M. Regler, *Data Analysis Techniques for High Energy Physics*, Cambridge University Press, 1990. See Chapter 1, "Real-time Data Triggering and Filtering."
3. Solenoidal Detector Collaboration, G. H. Trilling, *et al.*, *Technical Design of a Detector to be Operated at the Superconducting Super Collider*, SDC-92-201 & SSCL-SR-1215, Dallas, TX, 1992. See in particular Chapter 8, "Electronics."
4. A. Boyarski, T. Glanzman, F. A. Harris & F. C. Porter, "Report of the Computing Group," *Proceedings of the Workshop on Physics and Detector Issues for a High Luminosity Asymmetric B-Factory at SLAC*, SLAC Report 373, Stanford, CA, 1991.
5. Solenoidal Detector Collaboration, G. H. Trilling, *et al.*, *Expression of Interest by the Solenoidal Detector Collaboration to Construct and Operate a Detector at the Superconducting Super Collider*, Berkeley, CA, 1990.
6. T. Bolton, D. D. Briggs, C. Hearty, W. R. Innes, A. J. Lankford, P. C. Rowson, P. Rubin, M. H. Shaevitz & W. J. Wisnewski, "Report of the Trigger and Data Acquisition Group," *Proceedings of the Workshop on Physics and Detector Issues for a High Luminosity Asymmetric B-Factory at SLAC*, SLAC Report 373, Stanford, CA, 1991.