10

Electronics

10.1 Overview

10.1.1 Introduction

The electronics system encompasses all front-end, trigger, and data acquisition electronics. The environment of the high-luminosity PEP-II machine poses certain unique challenges to this system. PEP-II beam crossings occur at a rate of 238 MHz, which, in terms of the response time of the electronics, is essentially continuous, unlike previous e^+e^- colliders. PEP-II may have severe backgrounds, producing high occupancies. These high occupancies present a significant burden on the transport and recording of data. The rate of all physics processes to be recorded at the design luminosity of $3\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is expected to be 100 Hz. The electronics system will also need to accommodate the rate and background implications of a luminosity upgrade to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. These factors have resulted in the adoption of an innovative system architecture that departs significantly from previous implementations at conventional e^+e^- experiments. In this architecture, the electronics systems are capable of collecting data from the detector in realtime, and processing and transporting the data in parallel. The data are extensively buffered to avoid deadtime losses while making trigger decisions. In addition, the detector electronics, trigger logic, and data acquisition system are capable of simultaneous trigger processing of multiple events.

As described in Chapter 12, the backgrounds expected in PEP-II include particles lost near the IP, beam-gas interactions, and cosmic rays. The dominant source of occupancy is predicted to be the lost particle background, which depends on the beam intensity and the quality of the vacuum in the machine in the region from 3 to 50 m from the IP. At the design pressure of 1 nTorr, simulations predict raw hit rates of several tens of megahertz per layer in the silicon vertex detector, 100 kHz per layer in the drift chamber, and 15 kHz per crystal in the calorimeter. The first device to become unusable as background increases is the silicon vertex detector. This will happen at approximately ten times the nominal rate. Both long-term (ten-year) radiation damage and very high occupancies in the inner layers provide limits to useful running conditions. The electronics system is fully operational under background conditions up to ten times the nominal rate. At higher background conditions,

System	Number of Channels
Vertex Detector	150,000
Muon System (IFR)	42,000
DIRC	13,400
Calorimeter	7,000
Drift Chamber	7,000
Aerogel (ATC)	144

Table 10-1. *Channel count for each detector system.*

the system degrades gracefully, albeit with increasing deadtime, allowing the background and detector performance to be studied.

The principal features of the electronics system are: a common architecture for all detector electronic front-ends; common trigger, timing, and control interfaces; a common data acquisition interface; and a common online software environment from the lowest level within the front-end signal processors to the highest level user interfaces and event processing farms. By generalizing the architecture for all systems down to the point of contact with signal digitization, the overall maintenance, software, and training needed to operate the system are minimized, and engineering can properly be focused on detector-specific requirements.

The *BaBar* detector has approximately 220,000 signal channels (summarized in Table 10-1) distributed among six types of detectors, delivering signals with shaping times varying from 1 ns to 3 μs. Since the accelerator operates essentially continuously, the fast trigger will be unable to identify the specific beam crossing associated with an event. In fact, the event time will be uncertain to hundreds of nanoseconds. Consequently, a requirement of the electronics systems is that they record the raw data in real time and extract the data within the time uncertainty window. Providing sufficient bandwidth to transmit all of the raw event data within these time windows would be overwhelming, so as much signal processing of the data as possible is performed prior to transmission. This requires storage of the data while filtering is being performed and while the pertinent data (features) are being extracted.

All detector systems have a common interface for event readout, fast control, trigger, and detector control. Yet, within this commonality, all detector front-ends have unique signal processing requirements. While it is not possible to devise a uniform interface to the trigger for each detector system or to provide a uniform implementation of the storage of data awaiting a Level 1 decision, the same architecture of data storage and control has been applied to the signal processing chains for all detector systems. A block diagram of the major components of the electronics is shown in Figure 10-1. The following subsections give overviews of the front-end electronics, trigger, and data acquisition systems.

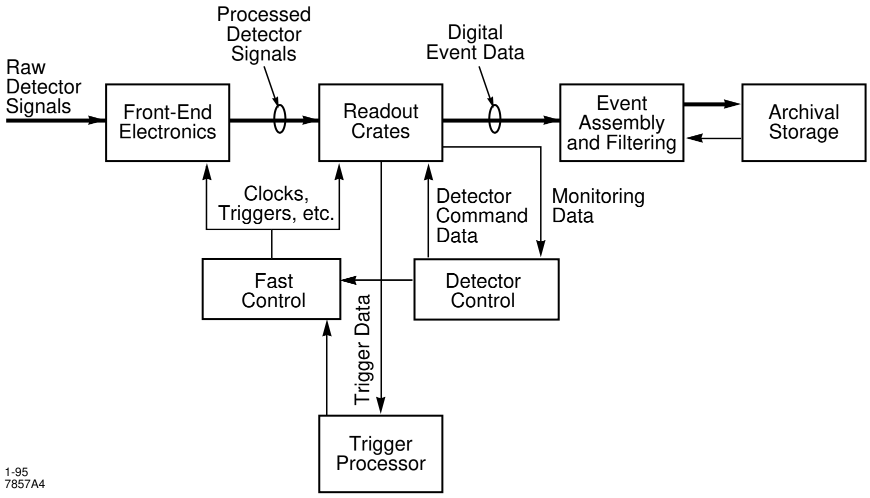

Figure 10-1. *Block diagram of the electronics system showing the principal communication paths. The dark arrows show the event data path.*

10.1.2 Front-End

In the *BaBar* nomenclature, the front-end is that portion of the electronics from the signal source on the detector through the digitizer. For those systems which have their digitizers mounted on the detector, the front-end also includes any additional detector-mounted parts. As shown in Figure 10-2, the front-ends are customized to fit the needs of the particular detector system. Figure 10-2(a) shows a generic front-end. The raw detector signals are appropriately processed and transmitted to the readout modules, which are housed in VME crates in the electronics house. The electronics house (in contrast to the detector) is accessible when the machine is running. Figure 10-2(b) depicts a front-end for the calorimeter. Since the event time is not known at this point, the digitization is done periodically, at a rate optimized for each system. Figure 10-2(c) is the front-end for the drift chamber. Here the digitization is done on the readout module. Figure 10-2(d) shows the front-end for the vertex detector, the instrumented flux return and the DIRC.

The processed detector signals are transmitted to the off-detector readout crates. For systems in which the first stage of buffering is on the readout modules, processed signals constitute a continuous data flow; selection of data by the triggers is accomplished on the readout modules. Systems which have buffers in the detector-mounted electronics send data selected by a Level 1 strobe only upon the request of the readout module. Examples of the first case are the drift chamber and the calorimeter. Examples of the second case are the vertex detector, the instrumented flux return and the DIRC.

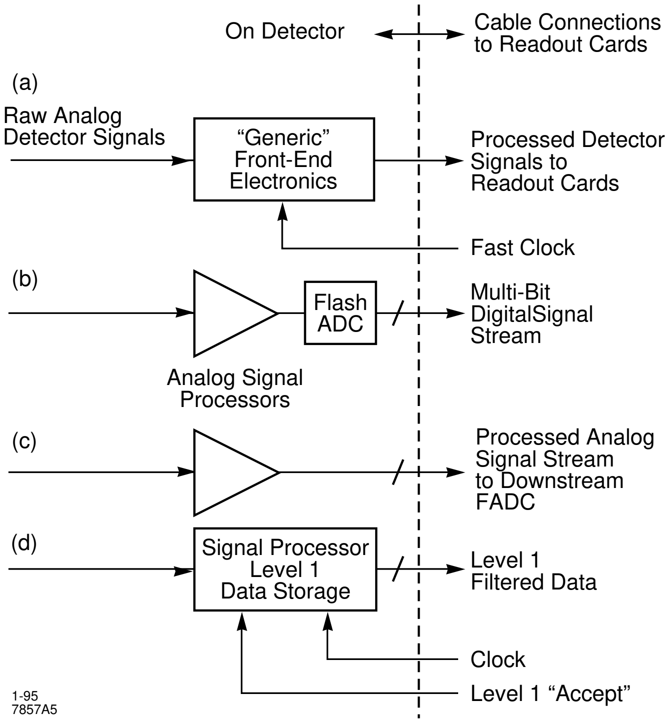

Figure 10-2. *Examples of front-end electronics for various detector technologies. Variations (a) through (d) are explained in the text.*

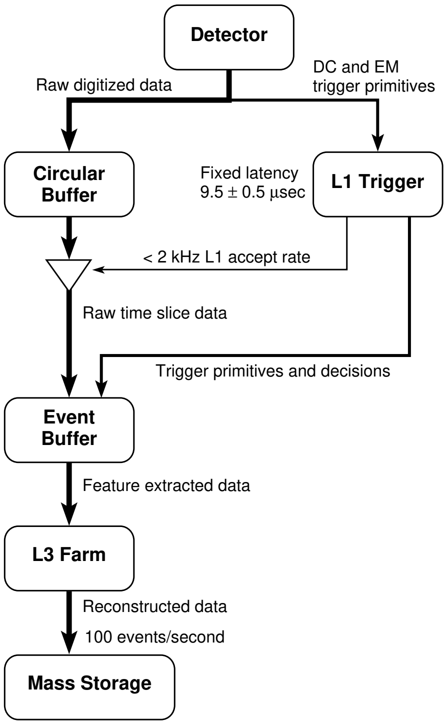

Figure 10-3. *Trigger structure and data acquisition data flow, showing levels and buffers. See text for a discussion of Level 2 options.*

10.1.3 Trigger

Definition of Trigger Levels and their Functional Requirements

The structure of the trigger and data flow are shown in Figure 10-3. The trigger:

- provides a front-end strobe (Level 1 strobe) after a latency between 9 and 10 μs as required by the finite depths of the front-end buffers (Chapter 4);
- initiates, in the readout modules, the extraction of hit times and magnitudes from the sample streams (feature extraction);
- initiates the collection of the data associated with an event in a single processor (event assembly); and

- selects a sample of events to be written to mass storage at a rate of no more than 100 Hz, as required by the online computing system.

This architecture is the simplest robust solution that performs the trigger functions, given the rate considerations.

The only practical method for delivering a Level 1 strobe in less than 10 μs is to use a hardware processor operating on a reduced representation of a subset of the data (trigger primitives). The most powerful method for reducing the trigger accept rate to 100 Hz with no loss to the physics program is to use the vertex detector to select events with multiple tracks originating from a narrowly defined primary vertex near the e^+e^- luminous region. Simulation studies show that the required vertex detector information cannot be made available to the Level 1 hardware trigger logic because the bandwidth required to supply sufficient position granularity to suppress background is comparable to the raw-data bandwidth required *after* the Level 1 strobe [Lev94c]. Because the vertex information cannot be included, the Level 1 accept rate will be considerably higher than the eventual rate of writing events to archival storage. Event assembly could proceed comfortably at a rate of 2.0 kHz, based on experience with other experiments [Boz93, Dou94, Buo, Cul91, Pat94]. Once the events have been assembled, there will be ample information with which to filter out the background events. Therefore, a two-level trigger containing a flexible Level 1 filter with a maximum output rate less than the maximum event assembly rate is necessary and sufficient.

To support the system architecture, dedicated data paths provide trigger data to the first level trigger hardware (Level 1). The maximum average acceptance rate for this trigger is required to be less than 2.0 kHz (Table 10-11 in Section 10.9). A final software filter (Level 3) running in commercial processors analyzes the event data from all detector systems and reduces the rate at which events are written to archival storage to less than 100 Hz. The trigger protocol is expected to be highly efficient for a variety of physics modes in addition to B physics events.

Each trigger level uses a unique set of data and algorithms, summarized in Table 10-2, that limits the accepted events. Successive trigger levels use increasing amounts of time to perform more intricate analyses of the data. Not all detectors contribute trigger data; those that do, preprocess it, *e.g.,* extract track segments from a collection of raw drift chamber hits. The results of this preprocessing are referred to generically as primitives. The primitives used by the Level 1 trigger are track segments from the drift chamber and energy sums from fixed sets of crystals in the calorimeter. The Level 1 trigger rate is dominated by the lost particle background with current estimates showing 2.0 kHz at ten times the nominal beam-pipe pressure. If needed, additional trigger criteria can reduce this rate with some small loss of non-B-physics efficiency.

Trigger	New Information	Algorithms
Level 1	DC Segments, EM Energy Sums	Track Counting, Energy Cuts, p_{t} Cuts, Matching
Level 2 or 3	Vertex System Occupancies	Primary Vertex Cuts
Level 3	Full Event Data	Final Event Selection

Table 10-2. *Trigger inputs and algorithms.*

Additional robustness is designed into the system via hooks for an additional Level 2 filter in case the actual Level 1 accept rate is higher than the maximum Level 3 event build rate, or if the feature extraction function requires more computing power and/or bandwidth than is currently estimated. The Level 2 trigger adds requirements on the track point-of-origin to reject non-annihilation backgrounds. It might also simplify the Level 1 trigger by doing the track transverse momentum cuts. Inputs to this trigger level would include all inputs to Level 1 and any additional signals needed as the algorithms are developed.

All trigger levels have the flexibility to examine event topology, count tracks, and match charged tracks to energy clusters. All levels have some capability in setting adjustable thresholds in energy or transverse momentum. This high degree of flexibility is required to reject known backgrounds under a range of conditions while maintaining very high efficiency for physics events.

10.1.4 Data Acquisition

The readout modules contain large memories that preserve the streams of digitized data. For systems with detector-mounted buffers, these memories store events after the Level 1 accept and are used to hold data that is ready for readout. The remaining systems also use this memory to buffer data during the Level 1 trigger latency.

Within the data acquisition system, there are four principal communication paths:

- An event data path by which all event data pass from the system electronics to the event collection system;
- A high speed clock, trigger, and control pathway for synchronizing the flow of data and controlling reset, initialization, and throttling of data;
- A relatively slow detector control interface between the system electronics and the online system; and

Detector System	Bytes per Event	Number of Sources	Bytes per Source per Event	Source Data Rates (MBytes/s)
Vertex	8400	5	1680	3.4
Drift Chamber	3600	18	216	0.4
DIRC	2700	1	2700	3.5
Aerogel	300	1	300	0.3
Calorimeter	7700	18	275	0.6
IFR	800	1	800	1.6
Trigger	800	3	400	0.8
Totals	24,300	47		

Table 10-3. *Summary of event data contributions from each detector type at ten times nominal occupancies and data rates per data source for each detector system for a 2.0 kHz trigger rate.*

- A trigger data path that conducts detector data to the trigger processor.

The detector control interface is responsible for initializing the system and for monitoring and debugging. Overall management, error logging, and databases are part of the online system described in Chapter 11.

The data acquisition interfaces are envisioned to be very similar across the systems in the following respects. The crates for each detector system will have

- readout modules,
- a fast control distribution module,
- a readout controller, and
- an event-data port card unit (data source).

For each event, the event-assembly network moves the approximately 25 kbytes of event data, summarized in Table 10-3, from the data sources to the Level 3 processor farm. The event-assembly network is implemented with commercial network technology. At a Level 1 trigger rate of 2.0 kHz, this technology is expected to be able to transport the approximately 50 Mbytes/s of data expected at 10 times the nominal design point. This type of technology is easily upgradable and can accommodate the needs of luminosity, trigger processing, and bandwidth upgrades.

10.2 Silicon Vertex Detector

This section describes the storage and transport of data from the silicon vertex detector transition card into the data acquisition system. The readout chip electronics and the transition card are described in Chapter 4.

10.2.1 Requirements

In the vertex detector readout, the data are digitized and sparsified at the detector front-end. Unlike the data from the drift chamber and calorimeter, the silicon vertex detector data must be buffered before being moved from the detector. This minimizes the deadtime that can arise from multiple triggers occurring on a timescale which is short compared with the readout time, or from events that generate large numbers of hits.

Because of the large number of signal channels in the silicon vertex detector (approximately 150,000), the electronics for this detector are required to produce sparsified data streams. For each hit that occurs within a specified time window (*e.g.,* $\pm 0.5\,\mu$s) of a Level 1 trigger, the system must generate three numbers: the number of the channel (strip) that was hit; a rough measure of the time-over-threshold of the pulse produced by the hit; and the time at which the pulse first crossed the threshold.

Data fragments consisting of the above three numbers are generated by the silicon vertex readout chips and placed into the chip's three-deep buffers. Subsequently, the data in these buffers are gathered event by event, and merged into larger data fragments, each covering a large number of hits associated with a particular trigger. Further numbers on particular performance parameters like dynamic range, noise levels, and speeds are given in Chapter 4.

10.2.2 Overview

A block diagram depicting the organization of the vertex detector electronics is shown in Figure 10-4. Each chip processes the signals from 128 strips as follows. A time-over-threshold signal is created, which is a logarithmic function of the pulse height. For each clock period, depending on whether the signal is above or below threshold, a binary digit with a value of **1** or **0**, respectively, is input to a latency buffer, awaiting Level 1 decisions. These bits are preserved for the latency period of the trigger, then erased. When a trigger is received, the time history of the stored bits is examined, and the number of **1**s within a certain period before and after the trigger is determined. This number represents a measure of the pulse height. In addition, the time with respect to the trigger of the start of the stored **1**s is

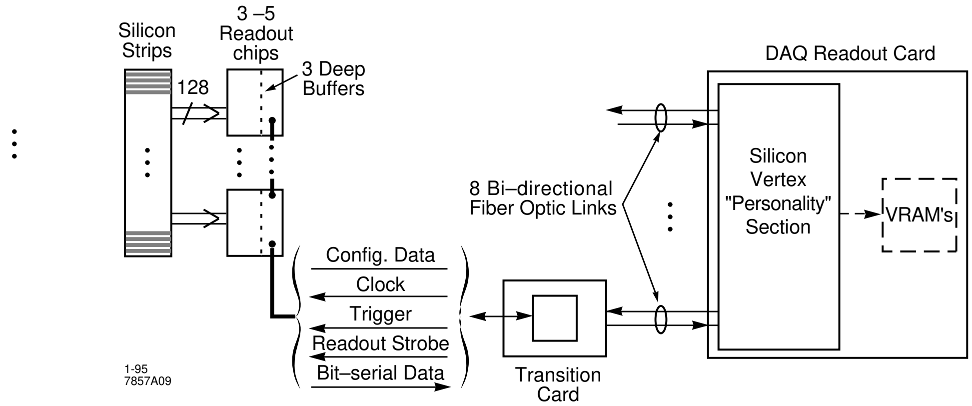

Figure 10-4. *Block diagram of the vertex detector electronics system.*

found. This represents a measure of the time of arrival of the pulse from the strip. These two numbers are then stored in the three-deep (FIFO) buffer for that channel. If there are no **1**s in the latency buffer close to the trigger, then a null-data indicator is stored in the three-deep buffer.

During the readout of an event, the data in the last levels of the buffers of all channels containing valid data are read out in bit-serial mode, and channel serial numbers are added. In this operation, several readout chips are connected in a daisy chain as shown in Figure 10-4. The data travel a short distance to a transition card, which launches the bit-serial stream on an optical fiber link destined for a readout module (Section 10.10.4).

The optical fiber link is bidirectional. In addition to the bit-serial data stream traveling toward the readout card, a set of signals is carried in the opposite direction by a second fiber. These signals include: a system clock, Level 1 triggers, event data readout strobes and controls, and several bits of configuration data. The configuration data can control certain functions of the readout chips such as putting a chip into calibration mode.

A single bundle of data lines plus control lines is assumed for each readout section. A readout section is either the z or ϕ side of a detector. To provide redundancy in case of a single point failure, pairs of readout sections (z and ϕ) on the same detector are grouped together such that the data and control lines of one side can service chips on both sides, if necessary. There are 208 readout sections servicing the silicon vertex detector.

The readout chips are daisy-chained on a serial communications pathway operating at 80 MHz. The data are passed through the chips rather than on a common bus. The advantage of this is that a chip can only break the line in one place as opposed to taking down an entire bus. The chips can pass their respective data in either direction on the pathway, and both ends of the pathway are accessible to external electronics.

The data transmission control is implemented as a redundant set of readout strobes. A pulse on the readout strobe gives the first chip in the daisy chain priority to begin sending its data. The packet of serialized data is always prefaced with a header and terminated by a trailer. When this data has been transmitted, the token giving priority to the data pathway is passed to the next chip in the daisy chain, and the process is repeated.

10.2.3 Data Format

The format of the bit-serial data travelling between the transition card and the readout card contains three fields: a header, a body, and a trailer. The header has a unique code that enables the personality card to locate unambiguously the start of an event data message. The body contains the information from a number of hits. The number of hits will, of course, vary from event to event. The information from each hit is contained in a subfield containing the three numbers—strip identification, pulse height, and time of arrival. These subfields are of constant length. The end of the subfields is announced by a unique trailer bit pattern.

10.2.4 Description of the Readout Module

As shown in Figure 10-5, the signals on eight optical fiber pairs are received by the personality module, which is physically part of the readout card. A generic readout card is shown in Figure 10-30 in Section 10.10.4. Here, the optical signals on the fiber are converted to bit-serial digital logic signals by the fiber optic receiver. Next, the serial-parallel converter converts the bit-serial data to word-serial data in which the words are 16 bits wide. The words are then passed to the link controller.

The link controller on the vertex personality module must identify the start of data from the header code and the end of data from the trailer code. Since each chip in the chain has included a trailer after its data, the data acquisition system can count how many chips sent data in response to a single read strobe. The system knows from a configuration database how many chips should report data from a given readout section, so it can recognize the end of the data for an event.

Relying on the format described above, the link controller in the personality module is able to extract the hit data from the incoming data stream and prepare it for forwarding to the VRAM on the DAQ readout card. When it senses the last trailer field, the link controller raises a “Done” flag. When all such flags are raised, an event-data-ready signal is strobed to signal the onboard CPU that a new event has arrived and to enable the issuance of the next readout strobe.

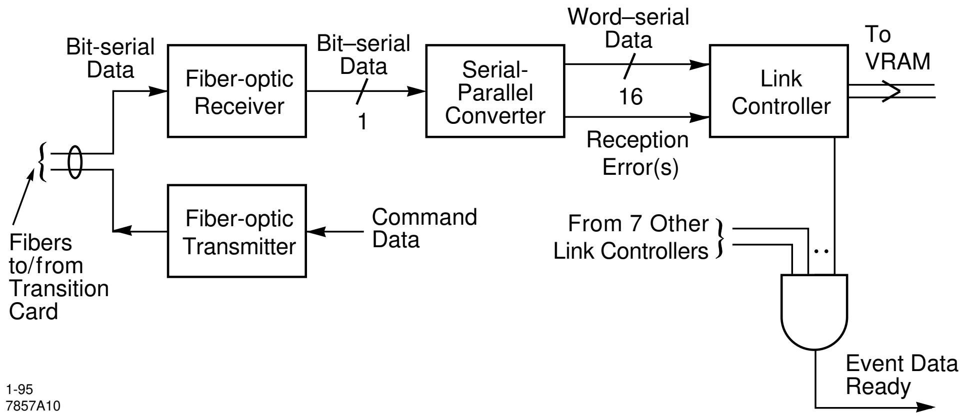

Figure 10-5. *Block diagram of the vertex detector specific section of the readout module.*

The readout strobe is not buffered on the chips. The DAQ readout card must never send a readout strobe until all of the data corresponding to the previous readout strobe have been read out.

The serial data stream enters the VRAM buffers synchronously with the 80 MHz serial transmission clock (5 MHz after 16 bit deserialization). The VRAM buffers are large enough to store approximately 500 events.

Each of the readout modules also has a processor for signal processing and message handling. Upon receipt of a Level 1 trigger, the processor copies all data from the relevant time slot of the VRAM to conventional RAM. The processor extracts the physical quantities (times and amplitudes) from the information stored and passes the result on for event assembly. Additional buffers are required to hold the post-processed events until they can be read out.

10.3 Drift Chamber

10.3.1 Requirements

The readout system for the drift chamber should not introduce any significant degradation of the performance of the chamber. The timing resolution should be sufficient to determine the distance of closest approach of a track to each wire with a precision better than the inherent chamber resolution. The charge deposition should be measured with an accuracy compatible with the statistical fluctuations in the formation of the primary clusters. Cell hit information should be made available for the segment finder logic with an efficiency approaching 100%

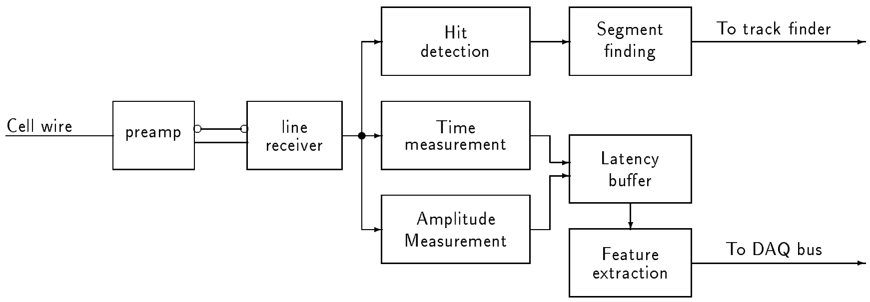

Figure 10-6. *Functional diagram of the drift chamber readout system.*

and with minimal contributions from electronic noise or neighboring channels. A functional diagram of the readout chain is shown in Figure 10-6.

The position measurement should have a precision of 140 μm. With an average drift velocity of about 25 μm/ns, this distance translates into a time precision of approximately 5 ns. A timing resolution of about 2 ns would then degrade the measurement by 10%.

The energy loss in the gas is usually measured by integrating the charge collected on a sense wire. The fluctuation in the number of primary clusters is about 7%, so the charge measurement requires a precision of a few percent.

10.3.2 Preamplifiers

The preamplifier is of the transimpedance type, giving an output voltage proportional to the instantaneous current flowing into the sense wire. The current shape from the drift chamber wire has a rapid rise when the charge multiplication occurs, and then falls inversely with time. For a shaping time of 6 ns and a gas gain of 2×10^4, the current pulse from one produced electron is about 60 nA, with a dynamic range of about 250. Thus, a noise figure of about 1000 electrons is required. The current cuts off when the last positive ion reaches a field wire 100 μs later. In the inner wires of the chamber, where the count rates might reach 80 kHz under $\times 10$ background conditions, there would be pileup on the tail of the previous event. The pileup would have little influence on the dE/dx information quality, but would cause some inefficiency in the discriminator. Therefore, the preamplifier includes a pole zero cancellation circuit with a time constant on the order of the drift period (300–500 ns). The preamplifier risetime will be optimized to the digitizing scheme selected.

The use of an unsaturated gas in the drift chamber requires careful control of the parameters that determine the gas density, *e.g.,* the gas temperature. The power consumption of the preamplifiers which are all connected to one end of the drift chamber is of concern in this regard. The preamplifier will be built as a bipolar semicustom ASIC, which allows a reduction in power of a factor of 10 compared to previous designs. A goal of 4 mW/channel is envisioned. The chip will have six channels, and several chips can be assembled on a preamplifier card to match the geometry of the drift chamber. Each preamplifier chip is individually protected by a polyfuse. In the event of a chip failure which shorts the power input at the chip, the polyfuse will open and allow the other preamplifier chips on the board to remain powered. The fuse can recover from a temporary short or a surge by cycling the power off and on. It should be noted that a loss of six cells on a layer does not prevent the segment finder from working.

10.3.3 Digitizer Options

Two classes of implementation schemes have been studied for the digitizer: those in which the time and amplitude measurements are done separately, and those in which the two measurements are combined in a single device. In the latter class, the flash analog-to-digital converter (FADC) is the natural candidate. It provides both time and amplitude measurement capability, does not require external gating, can digitize continuously with no deadtime, and is naturally compatible with a clocked dual-ported latency buffer. When operated synchronously at a submultiple of the machine frequency, the FADC provides an automatic rejection of the machine-related background noise. An FADC readout system can be assembled from off the shelf components from multiple manufacturers. An FADC-based readout system is also a formidable debugging tool—it is essentially a 7000-channel digital storage oscilloscope.

The second class of implementation schemes uses a specialized device to measure the time. For deadtimeless operation, a flash time to digital converter (FTDC), or time memory cell, will fulfill the requirements. The amplitude measurement is still achieved with an FADC but no time information is expected to be extracted from the data; thus the timing objectives can be achieved at a substantially lower clock rate. It is slightly more complex, not immune to machine related noise, and requires a low-time-walk discriminator to trigger the FTDC. In the all-FADC approach, a discriminator is also present to generate the hit information for the segment finder logic, but then the timing requirements are very modest.

An interesting variation of the FTDC/FADC approach is to build many such channels on a single ASIC, including a shallow latency buffer. The advantages are lower power dissipation, lower cost per channel for the components, and a considerable saving of space required on the readout boards. A larger number of channels can then be packed on every readout board, thus reducing the number of readout crates.

10.3.4 Implementing the All-FADC Scheme

An FADC readout system has been studied on the test drift chamber together with various algorithms to extract the time information. The sampling frequency for the tests was 250 MHz, close to the 238 MHz machine frequency. The performance at lower frequencies (250/2 or 250/3 MHz) could be deduced from the same measurements. Since commercial FADCs have a very short time aperture, it is sufficient to use one sample out of every two or three to simulate the operation at reduced clock rate. At 83 MHz(250/3), the position resolution achieved with the FADC readout channel was 135–142 μm. That corresponds to a degradation of 20% compared to the 115 μm achieved with the TDC readout. At 125 MHz, the degradation reduces to 10%. It should be noted that the above figures are somewhat unfair to the FADC method. The time-to-distance function is obtained by an iterative process that uses only the TDC time measurement information. The same function is applied to the FADC time measurement, though it has been optimized for the TDC system. Expected improvements in the time extraction algorithms may lead to the firm conclusion that one-third of the machine frequency (79 MHz) is a satisfactory sampling rate. But at this time, the all-FADC system operating at 119 MHz (one-half machine frequency) remains our baseline design. No technical uncertainty is involved in building such a system, and it has been demonstrated to achieve acceptable performance with a drift chamber very similar to the final one. FADCs which operate at 119 MHz are expensive and power hungry. The use of pairs of slower FADCs with interleaved sampling is being pursued. The uncertainties of this approach are discussed in Section 10.3.6.

Time Measurement

One of the difficulties in achieving a very good timing resolution with an FADC in the *BaBar* drift chamber originates from the low-pressure, low-Z gas. This choice minimizes the multiple scattering and yields very well-defined track trajectories. However, it also reduces the density of primary electron clusters along the track to about one cluster per millimeter. For tracks that directly cross a wire, this represents a fluctuation in drift time of about 15 ns for the arrival of the first cluster. Fortunately, this time difference reduces rapidly as the distance of the track from the wire increases and soon becomes a marginal effect. However, the coarse cluster granularity has a serious effect on the pulse shape of the signal extracted from the wire. It causes a succession of current pulses that can cover the full drift period in the case of an anode-crossing track. As a result, the pulse shapes are very different from event to event, making the time extraction a difficult task.

Various de-convolution methods have been tried. They succeed in separating the individual clusters when they are sufficiently far apart in time, but at the leading edge, where the geometrical effect makes the clusters concentrate for tracks away from the sense wire, the situation is very difficult. So far, the so-called TDC-mimic algorithm, which is a linear extrapolation of the first two bins to the baseline, gives the best results. Other algorithms are being studied. The timing performance might also benefit from tuning the preamplifier shaping. New, faster preamplifiers with adjustable shaping are now being installed on the test chamber to improve these measurements.

For the extrapolation algorithm to work accurately with small signals, it is necessary that the smallest signals, corresponding to single clusters, have an amplitude on the order of 10 ADC counts. This partially determines the dynamic range of the FADC. To prevent the digitizer from overflowing with large pulses, other multiplying factors have to be taken into account: 2 for track angle, 2.5 for highly ionizing particles, and 4 or 5 for cluster pileup, for a total factor of 250, or 8 bits. An effective dynamic range of 9 bits would be more comfortable.

Charge Collection

The evaluation of the primary charges deposited in the volume of a cell is made by summing the FADC samples over the drift interval. The statistical fluctuations in the formation of the primary clusters already introduce an average variation of 7% along the full length of the track. Thus, a measurement of the integrated charge with a precision of a few percent will not degrade the dE/dx evaluation, and a precision of 6 bits is sufficient for the charge measurement. If this requirement is combined with that of time evaluation, the FADCs should have an effective dynamic range of 8 to 9 bits and an accuracy of 6 bits. It would then be satisfactory in principle to use a 6-bit digitizer, and obtain the 8- to 9-bit dynamic range with a nonlinear or a multislope conversion.

Readout Module

The readout module is composed of two distinct sections: a front-end specific to the drift chamberand a common data acquisition section. The common section contains the latency buffers, some local RAM, the processor that will extract the physical features from the raw data, and an interface to the VME crate bus. The front-end section carries the FADC chips, the signal conditioning electronics, the discriminators that provide the wire hit information, and the segment finder logic, as shown in Figure 10-7. Each differential signal line from the preamplifier arrives at the readout board at a line receiver featuring a large common mode rejection. The output of the line receiver feeds a discriminator and a driver amplifier for the FADC chip input. This driver also includes a low pass filter to optimize the signal shape for the time extraction. An offset voltage is supplied to the driver from a 6-bit CMOS DAC, in order to adjust the FADC pedestal.

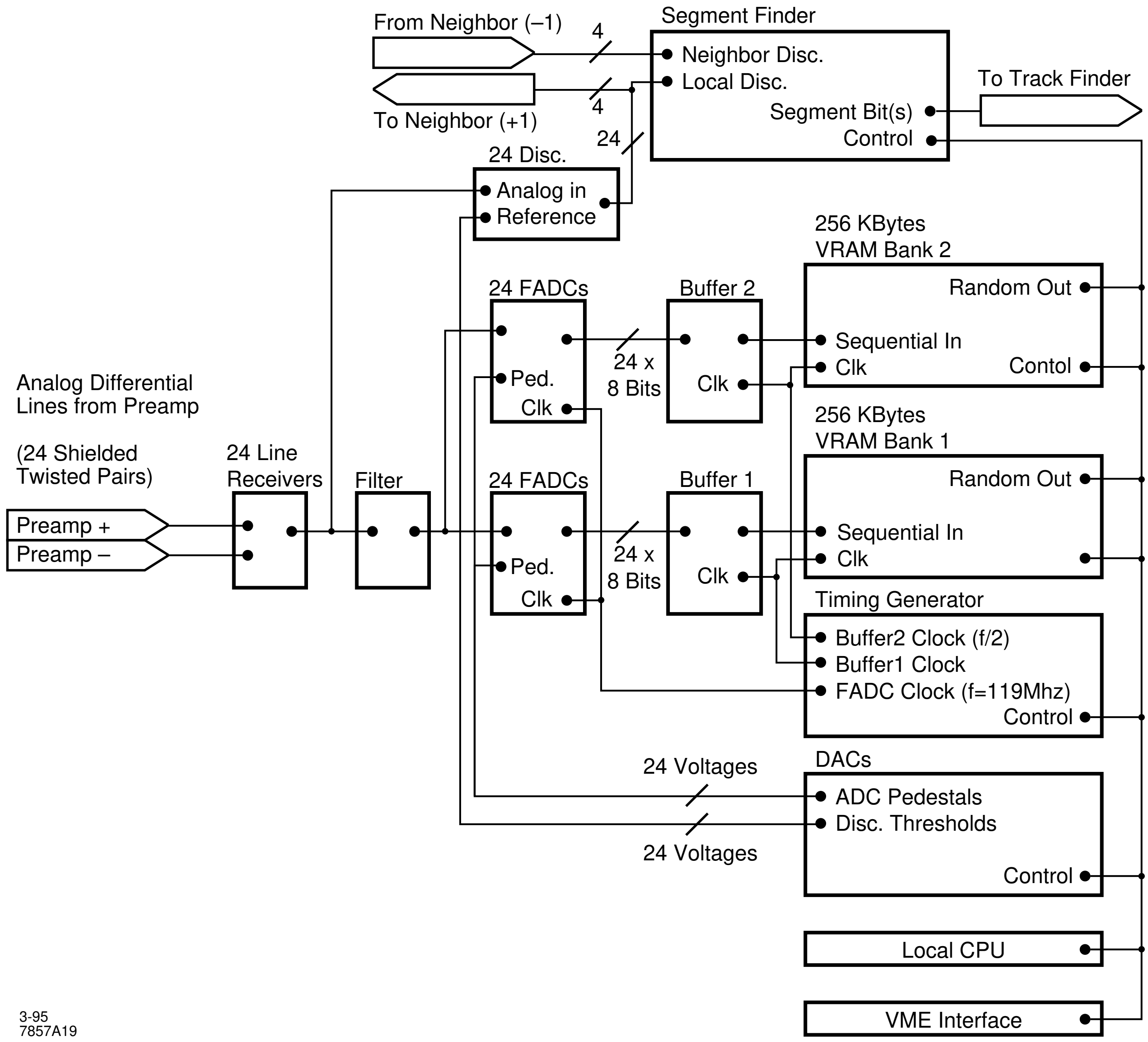

Figure 10-7. *Drift chamber module block diagram.*

The FADC chips are 8-bit devices operated at a clock frequency of 119 MHz, which is almost twice the maximum speed at which the VRAM buffer memories operate. Thus, it is necessary to use two interleaved VRAM units for each FADC, interfaced through two buffer latches. The chamber geometry will determine how many FADC/discriminator channels are put onto each board. For the baseline design, the chamber is organized into sets of four layers each, called superlayers, which are themselves organized into groups of 24 cells called supercells. Since the data acquisition board supplies the information to the segment finder, and the segment finder needs the information from one supercell, the number of channels on the board will be equal to, or a multiple of, the number of cells in a supercell. The segment finder logic also requires that some overlap be provided between adjacent supercells. An interconnect is thus required between neighboring readout modules, forming a daisy-chained closed loop. In the baseline configuration, three signals would be carried from one module to the next.

10.3.5 Calibration

The gain calibration is accomplished by injecting pulses at the preamplifier input. To limit the number of pulser cables going to the preamplifiers, a compromise is made between the flexibility in the trigger patterns and the size of the system. Each channel of a preamplifier integrated circuit can be pulsed independently. It is therefore possible to detect channels shorted together and to measure crosstalk. Each preamp cable has two or four pulser lines, for a total of about 1400 lines in the latter case. The pulser signals originate from the readout module. A lesser number of pulser lines is required if the pulses are distributed in a ring geometry. One pulser line is then connected to one of the four channels of each preamplifier chip belonging to that layer. This requires 320 pulser lines. Every line serves at least 20 channels, but this granularity is still small enough to generate useful patterns to exercise the trigger logic and the data acquisition system, in addition to the basic gain calibration function. The pulser system only calibrates the electronic gain. It must be combined with the information extracted from the data itself, such as the histograms of the integrated charge for each cell, in order to account for the wire gain.

10.3.6 Research and Development

Operation at a Lower Sampling Frequency

The research program concerning the FADC time extraction algorithms will continue, since even a slight improvement would justify reducing the sampling rate from 119 MHz to 79 MHz. At this lower rate, a CMOS FADC may be used, reducing cost by a factor of 5, power

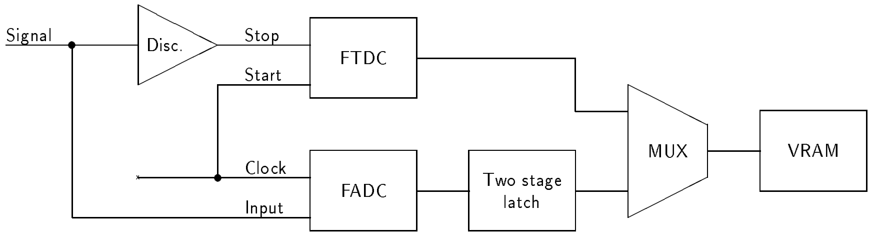

Figure 10-8. *Flash TDC/FADC channel principle.*

consumption by a factor of 4, and space on the the readout module by a factor of 10. Current candidate CMOS FADCs have only 6 bits and thus would have to be operated in a nonlinear mode. This has been done for other technologies but needs to be studied for CMOS devices. A test board will be constructed to evaluate this possibility and measure the maximum practical dynamic range expansion of the 6-bit CMOS FADC that is available commercially. The LBL group also has designed a 6-bit 100 MHz CMOS FADC prototype chip. It will be available for testing shortly. An 8-bit device could in principle be built with the same technology, with a somewhat lower maximum sampling frequency yet to be specified. It is also possible to use two slower FADC chips and operate them with split phase clocks to increase the sampling rate. Two CMOS units could be operated at 59.5 MHz to achieve an effective 119 MHz sampling rate. However, this requires a good matching of the transfer functions of the two FADC chips and might require a more expensive, more linear version of the chip. A test board will be made with CMOS devices to check the dynamic range compression method and the split phase technique.

Time Measurement with an FTDC

An alternate approach, aimed at lowering the sampling frequency, is to complement the FADC readout with an FTDC time measurement. The FTDC measures the time with very good resolution, and the FADC measures the charge only, while still providing a redundant lower precision time measurement. This approach is interesting if the sampling frequency is brought down to that of the VRAM buffers (40–60 MHz). Only one VRAM chip would then be required per channel. This is quite feasible since the FADC data would not be the primary source of time information.

As shown in Figure 10-8, the FTDC produces an encoded measurement of the elapsed time between a pulse of the sampling clock and the instant the discriminator fires. There is only one byte of TDC information for each hit, and it can be inserted in front of the FADC byte stream. The two latch stages following the FADC provide the required digital delay for that operation. This approach requires only one VRAM per channel instead of two, but the FTDC is an extra device. The FTDC function could also be achieved coarsely with a 6-bit CMOS FADC sampling a voltage ramp synchronized with the sampling frequency. In either case, the cost of this approach would be equivalent to that of the 79 MHz pure FADC operation and be subject to the same uncertainties relative to the dynamic range compression feasibility. However, since the FADC would no longer measure the time, the dynamic range requirement could probably be relaxed by a factor of 2.

Readout ASIC

An integrated drift chamber readout circuit could reduce the cost and complexity of the readout system. This readout IC operates like that for the silicon vertex detector, *i.e.,* multiple channels of electronics are read out through a single circuit only after a Level 1 accept. This architecture reduces the processing time on a DAQ card since only channels with hits are transferred from the readout IC. The architecture also minimizes the number of readout (VRAM) channels required on the DAQ cards and reduces the total number of readout crates required. Eight channels of drift chamber readout electronics are expected to fit on one integrated circuit.

The proposed drift chamber readout IC is composed of four main components: a time-to-digital converter (TDC) marks the time of the threshold crossing to an accuracy of less than 1 ns; a 6-bit FADC samples the drift chamber output pulse at 100 MHz, a dual-port DRAM stores the digitized drift chamber output pulse for the entire trigger latency of 10 μs; and a SRAM secondary buffer stores up to 3 μs-long sets of drift chamber hits.

The TDC is crucial in determining the exact time of the threshold crossing. The TDC operates by dividing the input clock period into 32 bins. The values of the bins are stored during each threshold crossing, and these values determine the portion of the clock period in which the threshold crossing occurred. The operation of the core of the TDC, the delay elements and the phase detector, has already been tested with a previous IC. An updated TDC currently in fabrication will test a new phase detector and decoding scheme.

An FADC built with auto-zeroed comparators has been designed. Simulations show that 6-bit operation at frequencies up to 100 MHz is feasible. The comparison scheme itself is good up to 10 bits. A 6-bit test circuit has been submitted for fabrication. Verification of operating parameters must be tested.

The DRAM and SRAM cores are designed and in fabrication. Some simple tests will be undertaken to determine if the memories operate correctly. These memories are built with memory generators. This allows adjustable size memories to be implemented for each design. Some slight modifications might be needed once the drift chamber IC channel pitch is determined. There should be no effect on the memory operation.

Once the above blocks have been tested, they can be connected at the board level to test the system functionality. A complete drift chamber channel can be designed with these blocks at the IC level concurrently with the system test. A four- or eight-channel prototype chip must be built to study power, noise, and crosstalk issues. The die size of an eight-channel chip is expected to be $16\,\mathrm{mm}^2$.

10.4 DIRC

10.4.1 Requirements

In the DIRC, the Cherenkov angle is determined from the position of hit photomultipliers among the array of 13,400 tubes situated on the standoff region backplane. The pulses are the result of the impacts of single photons on the photocathodes. The times of arrival of the photons from a given event are scattered within a 50 ns time window.

The electronics requirements are: to record single photoelectron hits quickly enough to avoid being swamped by the accidental rates, to provide 1 ns time resolution, to discriminate against background, to implement modest amplitude measurement (6 bits) in order to monitor the gain of the tubes and maintain the threshold settings, and to provide adequate bandwidth to safely transfer the physics data to the data acquisition system. Special calibration runs are needed to determine pedestals and gains.

Possible Readout Implementations

Several options have been considered, and their cost and performance evaluated: (1) full sampling, (2) full sampling with zero-suppression, (3) zero-suppression with local buffering with the Level 1 strobe starting the transmission of the data that arrived during the trigger latency, (4) the same as (3) but with only the data which arrived during the trigger timing window being transmitted. The above architectures can be implemented many ways. They differ in location of the components (on the detector or on the readout modules), the cable plant, number of readout modules and crates, amount of VRAM, and bandwidth of the transmission medium between the detector and the readout modules. A comparison of architectures is shown in Table 10-4.

At this time, options (3) and (4) are preferred because the number of cables is kept to a minimum, the amount of data to be sent to the data acquisition system is sufficiently reduced so that the whole DIRC readout resides in one or two crates, and custom integrated circuits can be used. The price is modest, and the custom ICs are not too complex.

Option	Number of Cables	On Standoff	Mbytes/s	PMs per VRAM Chip	VRAM Chips	Readout Modules	Readout Crates
1	13.4k	No	215k	2	6700	210	14
2	—	Yes	6k	27	500	65	4
3	—	Yes	100	1.34k	10	3	1
4	—	Yes	10	13.4k	1	1	1

Table 10-4. *Comparison of four DIRC readout options.*

Choosing between options (3) and (4) requires additional detailed study to judge whether the gain in bandwidth justifies the extra complexity of the digital front-end logic, since these added features can be easily implemented on the readout module.

Although the proposed architectures make the use of custom chips attractive, it has been decided to present, as the baseline implementation, a system in which discrete components are used wherever possible. It is believed, however, that an optimal system will be built mostly out of custom or semicustom integrated circuits for lower cost, improved performance, and more reliability. Such an implementation is explored in Section 10.4.3. The baseline solution implements option (4) as far as the functionalities are concerned, but the bandwidth characteristics are derived assuming option (3). This allows option (3) to be a fall-back solution for option (4).

10.4.2 Baseline Design

Overview

The plan is to install the front-end electronics as close as possible to the tubes. A data driven scheme has been devised in which digitization occurs only when the pulse is above a threshold. The time of each hit is measured to 1 ns with higher precision left as an option. Amplitude measurements with 6-bit precision are foreseen for calibration events only. Although the pulse height carries no physical information on a single hit, it would be useful to record it on an event-by-event basis to sort out pileups and provide assistance with debugging. This possibility is further explored in Section 10.4.3. An option is to add a high threshold discriminator on each channel to flag hits with anomalously high pulse heights.

Digitized data are stored for the duration of the Level 1 accept latency (10 μs) and are read out upon receipt of a Level 1 strobe. An intermediate stage between the front-end and the readout module is needed to synchronize the data flow and multiplex it to 15 optical links. One readout crate houses the entire DIRC readout.

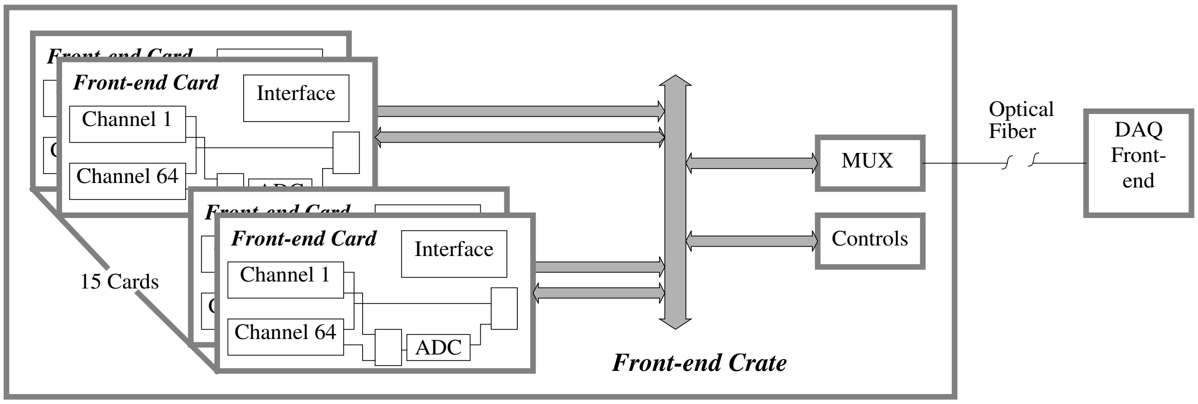

Figure 10-9. *Layout of the DIRC front-end electronics.*

The layout (Figure 10-9) calls for circuit boards located in crates situated near the tubes. Each card would carry the front-end electronics for 64 tubes. The idea of integrating the front-end electronics on boards shared with the high voltage system, attached to the back of the tubes, has been abandoned for the present because of the complexity of the interconnections and because the front-end electronics would get buried within the magnetic shield of the standoff box.

Short cables connect the bases to the circuit boards. A block diagram of the front-end is given in Figure 10-10. The main features of each channel are: timing—a constant fraction discriminator (CFD), TDC, and FIFO; the readout logic; and the slow analog calibration chain. A preamplifier with a gain of four and a bandwidth of 1 GHz may be needed at the input, depending on the choice of photomultiplier tube. Most of these features can be implemented with off-the-shelf components. A more detailed description of the items is given below.

Timing Chain

After the CFD, the time is measured by a double-tiered TDC. A coarse range is obtained by parallel loading 10 bits of a time register from the time counter, which records one pulse out of 16 from the storage ring rf (1 tick every 33.6 ns). Time measurement over the total Level 1 latency requires 14 bits. Sixteen bits have been allowed. The vernier range is obtained by using a high-precision 6-bit TDC. Early versions of this device have been built for the DELPHI experiment [Gen93]. The output of the 16-bit register is connected to a FIFO deep enough to keep 16 hits. A time window comparator and interface, and readout

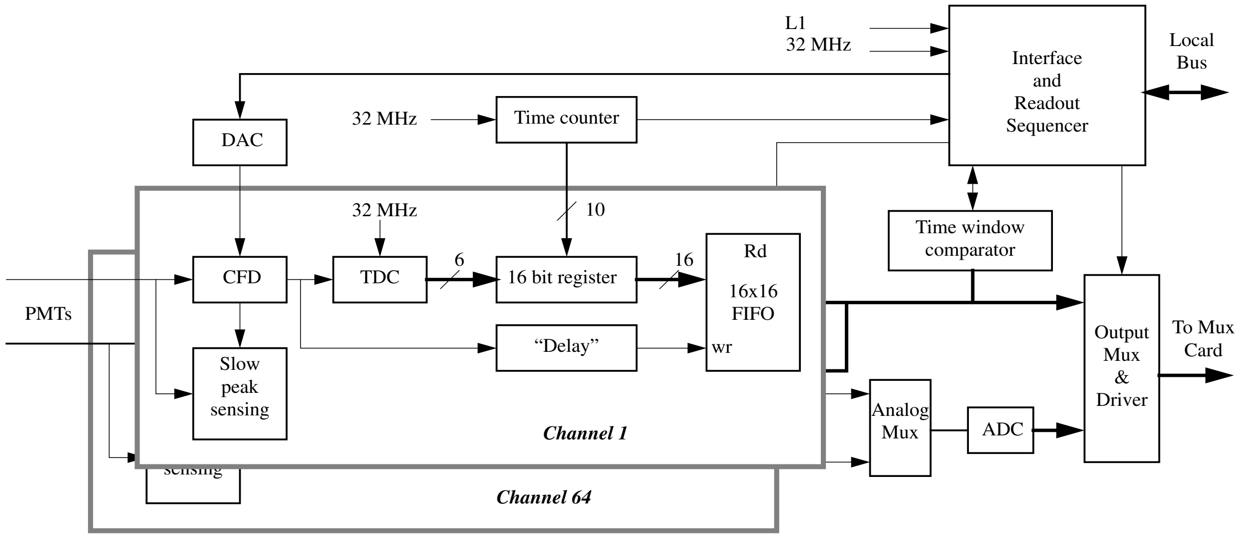

Figure 10-10. *Block diagram of the DIRC front-end electronics.*

sequencer enable the transmission of the correct data. The trigger time is compared to the hit times stored in the FIFO, and those which are in time are transmitted. The TDC and the FIFOs will be implemented inside a 64-input standard cell CMOS digital chip. The interface and readout sequencer adds the address of the photomultiplier within the card to the data word, which becomes a three-byte string. The output MUX/driver sends the data to the multiplexer boards.

Analog Chain

The pulse height will be measured during calibration. This requires an analog chain incorporating slow, charge-sensitive circuitry on each tube, which feeds a 6-bit ADC through a 16-to-1 analog multiplexer.

Multiplexing

From each front-end circuit board, the three bytes of data are sent at 400 kbytes/s to an intermediate crate. Assuming 20 Mbytes/s for the bandwidth of the readout crate backplane bus and a safety factor of about 4 to absorb the fluctuations in event size and trigger timing leads to a layout with 15 circuit boards per front-end crate. To minimize the cable plant, it is advantageous to add a multiplexer board in the sixteenth slot of such a crate. Each multiplexer board will concentrate the data of its front-end crate onto a 12 Mbytes/s optical

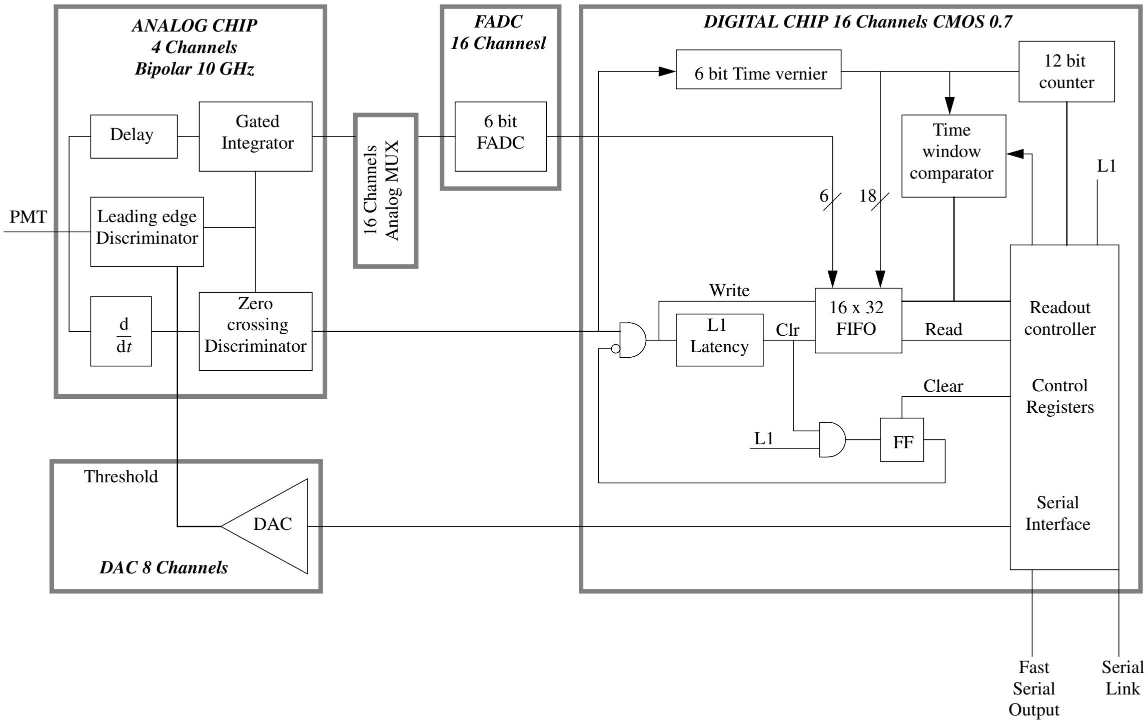

Figure 10-11. *Block diagram of the DIRC front-end electronics showing the suggested implementation with ICs.*

fiber serial link. This board will also add a 4-bit board address to make the data word 4 bytes. There are 16 front-end crates, each with its own readout fiber.

10.4.3 Research and Development

Most of the front-end described above is assembled from commercial components. It is well understood and buildable and has been used primarily for budgeting purposes. Only the TDC and the FIFOs are ICs. It is believed that a somewhat improved architecture can be built by using ICs throughout (Figure 10-11). Although such a design entails more development work, it should cost less and lead to a more reliable system.

In this scheme, the basic granularity is 16, although a better value may emerge from simulations. The philosophy of the timing measurement in this implementation is exactly the same as in the baseline design. The difference is that the amplitude will be measured during data taking. There is one FADC (6-bit, 100 MHz) serving 16 photomultipliers. If the rate is too high, not all tubes from a given readout section of 16 in an event will have their amplitudes measured.

The most elegant layout calls for circuit boards, attached behind the tube bases, which carry the front-end electronics for 16 tubes. There is enough room on these boards to consider sharing them with the high voltage system, should we choose to do so. Short cables connect the bases to the circuit boards. On each PC board, there are four analog chips, one or two digital chips, an analog multiplexer, a 6-bit 100 MHz FADC, two eight-fold DACs, and the output circuitry to send the data to the readout modules via digital multiplexers. More detailed descriptions of these are given below.

Analog Chip

An analog chip services four photomultipliers. For each tube, there is a discriminator and a shaping circuit to feed a charge-sensitive ADC. From the analog chip emerge the four discriminated output pulses, which are fed to the digital chip, and the four analog outputs of the shaper, which are directed to a 16-input/1-output analog multiplexer. In addition to the four signal inputs, the analog chip receives four threshold inputs from DACs. The choice among bipolar, CMOS, and BiCMOS processes is being studied.

Digital Chip

Each circuit board has one digital chip which accepts inputs from 16 tubes. The time measurement is identical to that of the baseline design. The interface and readout sequencer and the time window comparator could be made as a second separate chip or remain discrete components. These decisions will be made as the design matures. The digital chip receives 16 signals from the analog chip, the master clock, the trigger signal, and a calibration line used to measure the pedestals. The calibration proceeds as follows. A pulser controlled by the data acquisition system generates a signal which enables the discriminator threshold to be bypassed. The signal is also delayed to provide a pedestal trigger with the proper latency. For calibration or noisy channel suppression, individual tubes can be masked using a slow serial command line on the circuit board. The outputs of the digital chip are a 5-byte data bus daisy-chained across six circuit boards towards the fiber driver, and a 4-bit address bus to command the analog multiplexer. Standard cell CMOS is the technology chosen for the digital chip.

Front-End Layout

Nearly one thousand circuit boards will be used, each having four analog chips, a digital chip, an analog multiplexer, an FADC, and two eight-fold DACs. On top of the input and output signals already mentioned, cabling will carry the low power distribution lines. Slow control will have to be implemented locally. These features have yet to be designed.

Data Transmission to the Readout Crates

The data rate from all the digital chips is 100 Mbytes/s. Assuming the front-end readout modules accept data at 12 MHz, the flow must be shared among ten streams, each receiving data from 84 front-end sections. This corresponds to ten multiplexers with 84 inputs and one output, where a single address byte, that of the readout section, will be added to the data. Optical fibers similar to those considered for the vertex detector are appropriate. Alternatively, two 1 Gbytes/s optical fibers can be used.

10.4.4 High Voltage System

If a traditional resistor-divider photomultiplier base is chosen for the DIRC photodetector, a photomultiplier with a maximum operating voltage of 2 kV will draw a current of approximately 0.3 mA or less. In the proposed system, the high voltage will be distributed over long cables to a HV motherboard mounted on the photomultiplier support structure. Each motherboard services 16 bases, for a total current load of approximately 5 mA per HV channel. The motherboards will be connected to commercial multichannel high-voltage power supplies.

The high-voltage power would be distributed via standard RG-59 HV coaxial cable, utilizing a new connector scheme developed by LeCroy in collaboration with the AMP corporation. This connector is an eight-fold coaxial block connector, fully rated and UL listed for 5000 V. This results in significant savings over traditional SHV coaxial connectors. Connections to the motherboards would be made using standard SHV connectors.

10.4.5 Low Voltage Power Supplies and Control Systems

The on-detector front-end electronics has been estimated to dissipate a total of about 20 kW. The standard crates in the baseline design provide the needed low voltage and are water cooled. Control systems have yet to be designed.

10.5 Aerogel Threshold Cherenkov Counter (ATC)

10.5.1 Requirements and Overview

The signal processing provided by the off-detector electronics should be capable of fully exploiting the intrinsic performance of the ATC. For good efficiency, the threshold should be set between 1 and 1.5 photoelectrons and the timing resolution should be less than a few tens of nanoseconds in order to reduce the background from off-beam particles showering into the detector. An approach based on discriminators and TDCs, or one based on flash ADCs, can be used to achieve this goal. In the framework of the *BaBar* pipelined DAQ system, the most natural choice is the second of these options. This also provides some additional advantages:

- greater flexibility in setting different thresholds depending on the expected signal amplitude, which is a function of the particle momentum;
- shorter possible deadtime in high background conditions; and
- much more powerful diagnostic capability.

In the baseline design of the ATC detector, fine-mesh photomultiplier tubes (FM PMT) are used as photon detectors. After a possible front-end amplification and shaping, these can have an average pulse height and a shaping time similar to those expected from the drift chamber preamplifiers. Therefore, it may be convenient to use the drift chamber readout modules. The dynamic range of the expected ATC may be easily accommodated in the 8-bit effective range of the drift chamber readout system.

10.5.2 Front-End Electronics

With the Hamamatsu fine mesh PMT available today, the gain in a magnetic field of 1.5 T is about 10^5 and a low noise current preamplifier is required. To obtain a signal suitable for a commercial ADC, a transresistance of $60\,\mathrm{mV}/\mu\mathrm{A}$ is sufficient. This would give a signal of $\sim$50 mV for a single photoelectron pulse at 10^5 PMT gain. This value is considered to be the maximum requirement, since the PMT response still can be optimized with respect to the orientation in the magnetic field and improvements are expected in the performance of the next generation of PMTs. Using the specifications of a typical amplifier like the LeCroy TRA1000, one obtains an estimate of the noise of 0.5 fC FWHM, which is equivalent to 0.03 photoelectrons. The preamplifier is to be mounted directly on the PC board that supports the PMT and can be connected to the off-detector electronics via twisted-pair shielded cables.

10.5.3 Readout

The details of the drift chamber readout board appear in Section 10.3. We plan to make use of the processing power available on the board to perform the feature extraction after the arrival of the first level trigger. Once converted into amplitude and time information, the amount of data to be shipped to the event builder would be negligible, on the order of a fraction of a kilobyte. Provisions will be made to optionally send the sampled waveforms for diagnostic and calibration purposes.

High Voltage

The fine mesh PMT will be operated at 2.5 kV, with a steady current of approximately 2 mA. Existing commercial modular high voltage power supply systems are well-suited to the ATC HV power requirements. The ATC requires 144 HV channels. In the proposed system, the high voltage will be distributed over 12 cables to the PC boards where they will be split into two 6-pin connectors, each carrying the HV power needed by a PC board.

10.6 Calorimeter

The electronics for the CsI(Tl) calorimeter must read out the 6780 crystals in the calorimeter and provide information to the lowest level trigger as well as be able to respond to a trigger, and fetch and process the data corresponding to that trigger. This section describes the requirements of the calorimeter electronics and a proposed architecture which meets those requirements [Hal94].

10.6.1 Requirements

The following items are the primary requirements [Req94] for the calorimeter electronics:

The calorimeter data acquisition system. The system should not make the resolution for any photon with $E_\gamma \geq 20\,\mathrm{MeV}$ significantly worse than that due to the intrinsic properties of the detector alone. Specifically, the data acquisition contribution to the error should be less than half that of the CsI.

Dynamic Range. The calorimeter should have a dynamic range capable of reading out showers in beam events with energies from 20 MeV to 13 GeV with a resolution of 100–200 keV at 20 MeV and also provide adequate resolution for reading out low-energy gamma rays from radioactive sources for calibration. The dynamic range required for beam events is 17 bits. Allowing for a source calibration requires 18 bits of dynamic range, which is the design goal.

Incoherent Noise. The rms incoherent noise for a single crystal should not exceed an equivalent noise charge (ENC) of 350 electrons. This is the best that can be done at acceptable cost. This number combined with the number of photoelectrons collected for energy deposited (Chapter 7) determines the electronics contribution to the shower energy resolution.

Digitization Contribution. The contribution to the resolution from the digitization should be less than half the CsI(Tl) resolution. This requirement can be met with a 10-bit ADC.

Reliability. The electronics components which are inaccessible must be provided in two independent sets. Except for the crystals in the endcaps, access to components mounted on the crystals will be impossible. Additional components are located inside the structure on the end of the detector, but will be accessible by opening the detector endcaps.

Response to Background. The electronics should be designed to provide optimal performance with the expected levels of background. At high background rates, soft photons provide a source of noise which must be dealt with in the calorimeter (Chapter 12). This requirement makes demands on the shaping time of the electronics and the signal processing that must be performed on the output of the calorimeter.

10.6.2 Overview

The dynamic range requirement of the calorimeter combined with the large potential for beam-related rf pickup at PEP-II make digitizing the signal close to the CsI(Tl) crystals advantageous. The proposed design minimizes the cable length and dynamic range of analog signals in the system, and places the digitization just outside the rf-shield of the calorimeter, allowing access to the ADCs.

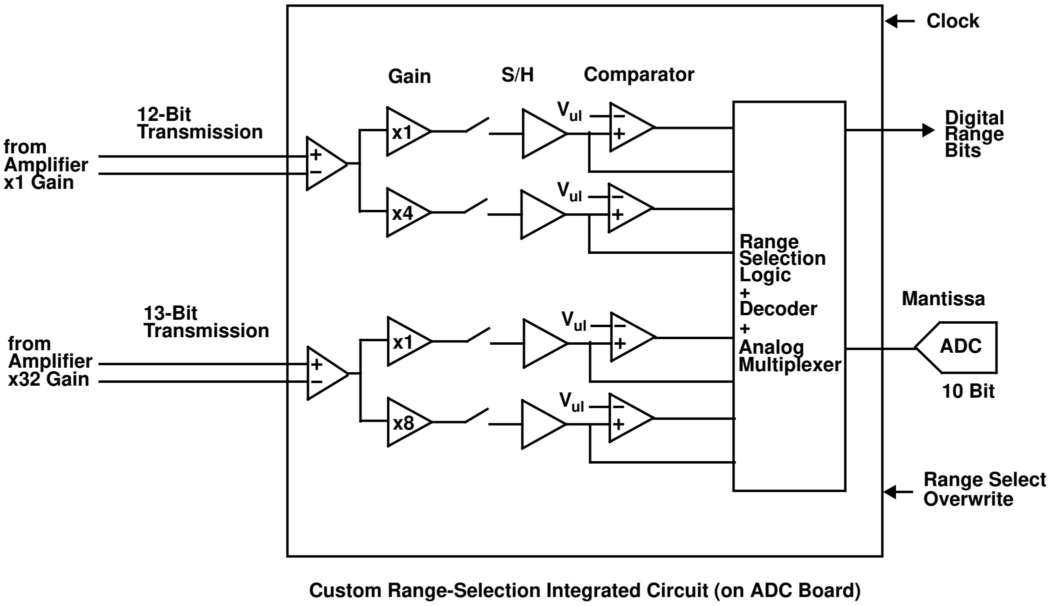

Figure 10-12. *Block diagram of gain stages and range encoding.*

To minimize the distance the 18-bit dynamic range analog signal travels, an amplifier with two gain stages is mounted within a shielded assembly on the rear of each crystal. By choosing amplifiers with gains of unity and 32, only a 13-bit dynamic range is required on the analog cables which go from the crystals to the digitizing board. At the digitizing board, a custom IC receives the two analog signals and provides two additional gain ranges. The outputs of this device are an analog signal, which is sent to an ADC at one of four possible overall amplifications, and two bits which specify which of the four gain ranges is being used. The additional gain stages allow an 18-bit dynamic range floating-point number to be constructed from a 10-bit digitization of the analog signal (the mantissa) and the two bits specifying the range (the exponent). Figure 10-12 shows a block diagram of the layout of the gain stages and range encoding.

Because it will not be possible to know which samples contain useful information until a trigger decision is made, the digitized data flow down two paths. The first path places the digitized floating-point data in a VRAM buffer where they wait for a trigger decision. When a trigger is received, the data acquisition system must use the trigger time to select the correct range of stored data to analyze and process and then transfer these data to the next level. In addition, the data from the calorimeter are used in the trigger decision, so digital sums of energies must be formed and sent continuously to the central calorimeter trigger logic. The trigger sums are more easily performed on the digital stream as the multiple gains stages make analog sums more complicated.

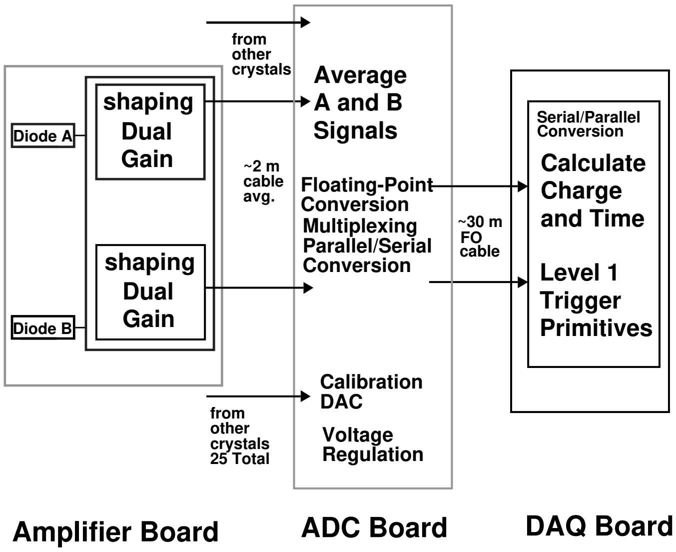

Figure 10-13. *Calorimeter system diagram showing preamp, ADC, and data acquisition (DAQ) functions.*

Figure 10-13 shows a block diagram of the components of the proposed readout system. The system has three main board types.

Preamp boards. One of these is mounted on the back of each CsI(Tl) crystal. They are not accessible after the calorimeter is constructed, so they must be reliable and contain a minimum of components. Each board provides an independent channel for each of the two photodiodes on the crystal (outputs A and B). Since it contains the first amplification stage, it must be well shielded. In addition, the power consumption should be low, as the temperature of the crystals needs to be regulated.

ADC boards. These are located at the end of the detector, just outside the rf-shield of the calorimeter. They should be accessible on the timescale of one day. Each board receives the signals from the A and B diodes on each of the approximately 25 crystals corresponding to the trigger towers. It contains the custom range encoding chip which chooses the gain then outputs two bits specifying the gain used. The board also includes the ADC chips, and its output is a serial digital stream.

Readout modules. These boards each receive the serial digital data from one ADC board, then de-serialize the data and store them in a VRAM buffer for later analysis. They also form trigger sums and ship the trigger data to the central calorimeter trigger logic. It is

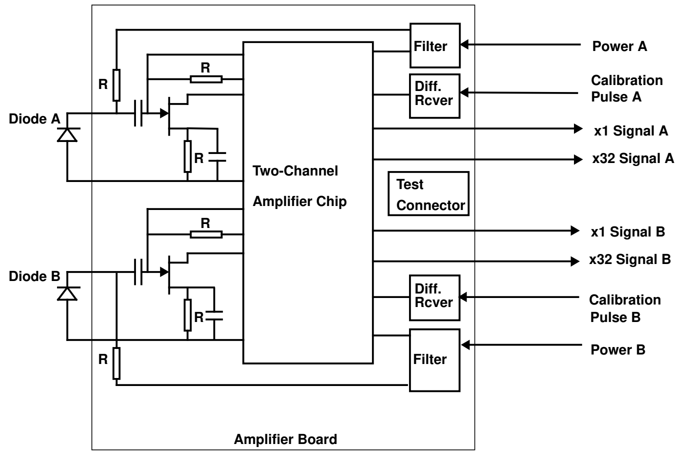

Figure 10-14. *Block diagram of preamplifier.*

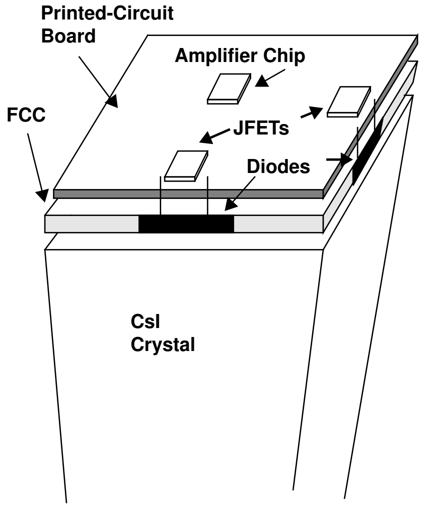

Figure 10-15. *Calorimeter preamp/crystal shield assembly.*

envisaged that there will be processing power on these boards to apply the gain and pedestal corrections to each channel and to extract the total charge and time information for each crystal when a trigger has been received.

Each of these components is discussed in more detail below.

10.6.3 The Preamplifier Card

Figure 10-14 shows a block diagram of the preamplifier board. It connects directly to the photodiodes on the CsI(Tl) crystals and provides, on output twisted pair cables, signals with $\times 1$ and $\times 32$ gain for both the A and B photodiodes. The main components are a custom two-channel amplifier chip (or perhaps two one-channel chips), and an external JFET. The shaping time for a CsI(Tl) pulse is optimally near $3\,\mu$s; however, the presence of background in the form of a large number of low-energy photons may mean that shorter shaping times are needed to give better background rejection. Shaping times less than $\sim 1\,\mu$s mean a significant loss of signal from the crystal. In the following, a shaping time of between 1 and $3\,\mu$s is assumed.

Providing most of the circuitry for the preamplifier on a single custom integrated circuit minimizes possible breakdown from faulty connections between components. The preamp IC contains completely separate paths for the power, signal, and calibration of both diodes. The only element in common between the two pathways is the substrate. The preamp chip provides charge-to-voltage conversion, shaping, and dual-gain differential driver stages at $\times 1$ and $\times 32$ gain.

The 18-bit dynamic range input to the preamp board requires that the diode and preamp be contained within a single shielded assembly. One possible layout of the crystal/diode/preamp assembly is shown in Figure 10-15. In this design, the preamp board is mounted on a stand attached to the crystal assembly. The diodes are then connected directly to the JFET on the preamp board, minimizing the distance over which the 18-bit analog signal passes. This crystal/diode/preamp package is enclosed in a metal shield, which can also be used for cooling. The heat output of the preamplifier board is estimated to be between 50 and 200 mW per channel.

Each preamp board requires many cables in order to provide independent power, calibration, and signal paths for both photodiodes. A reasonable estimate of the cable plant is:

- Power: V_{dd}, bias for the photodiode, ground, V_{FET};
- Signal: twisted pair for $\times 1$ and $\times 32$ signals; and
- Calibration: differential calibration pulse.

Thus, there are ten wires for each diode. Separate cables are provided for the A and B diodes. These cables need to be approximately 2 m in length. Filtering for the power lines and the receivers for the calibration pulses are provided on the preamp board.

10.6.4 Digitizing Board

The digitizing board collects the analog signals from groups of preamp boards, digitizes the data, and ships the digital data in serial form to the readout boards. Approximately 280 of these are needed to read out the A and B diodes from all the calorimeter crystals. Figure 10-16 shows a block diagram of the ADC board. A working assumption is that each board serves both diodes from 25 preamplifier boards. This board must provide the following basic functions:

- For each input channel, it must provide a floating-point number covering an 18-bit dynamic range. Our solution is to have a 10-bit digital mantissa and 2-bit exponent. This requires a 10-bit ADC.
- It must sample with a period which is short compared to the shaping time after digital filtering. This is needed to reduce background and improve the feature extraction. A reasonable maximum rate at which the ADC board will be required to digitize a single channel is 4 MHz (five samples over a peak for a 1.25 μs shaping time).

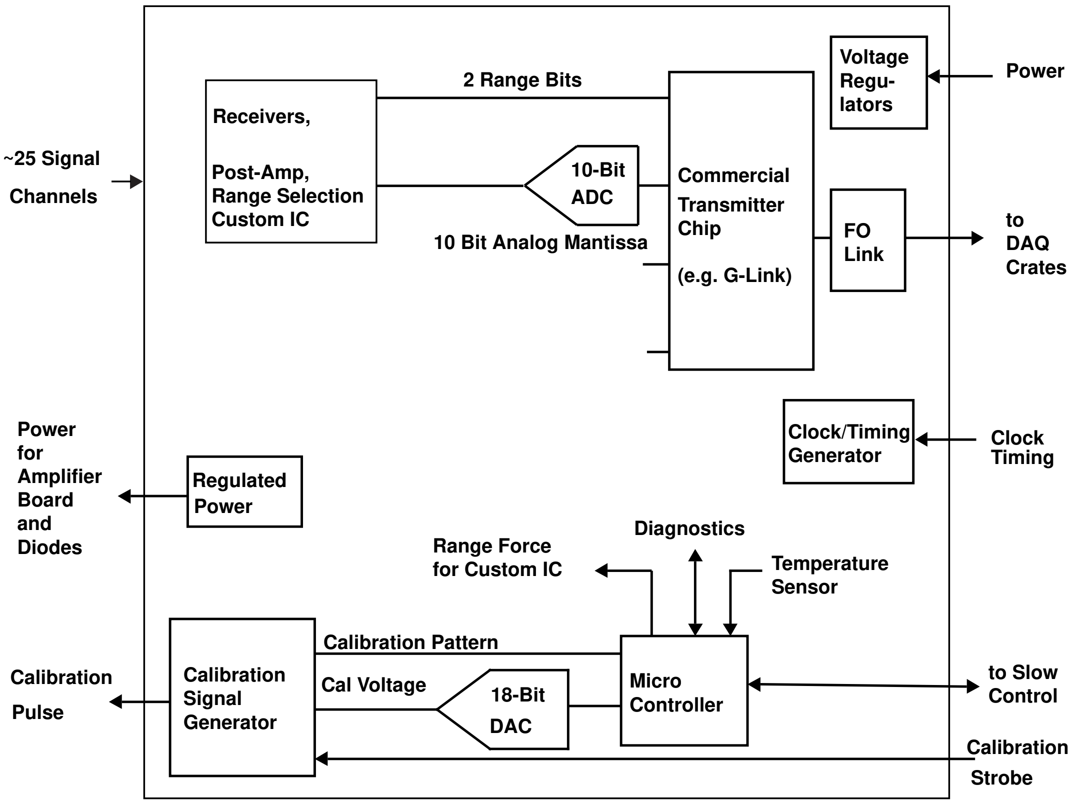

Figure 10-16. *Block diagram of ADC board functions.*

- It must provide the calibration signals to the preamp boards. In addition, the board should be able to control the range encoding chip and perform diagnostics and monitoring on the preamp and ADC boards.
- It must provide the power to the preamp boards.

One of the key components of the ADC board is the custom analog range encoding (CARE) chip which takes the four analog input signals from the A and B diodes ($\times 1$ and $\times 32$ for each of A and B), averages the A and B signals on the appropriate range, and then passes these through two more gain stages. The outputs of these four gain stages are sampled at the digitization sampling frequency and are compared to threshold levels. The chip places on its output the analog signal associated with the first gain stage not above threshold (for example 90% of the full range for that gain stage) and two bits signifying which of the four gain stages were used. The CARE chip can select one of three modes for operation:

- use input from only the A diode;

Energy Range (MeV)	Bin Size (MeV)	RMS Resolution (MeV)
0 to 46	0.050	0.029
46 to 366	0.40	0.229
366 to 2925	3.17	1.83
2925 to 13000	12.69	7.33

Table 10-5. *Calorimeter bin size and resolution as a function of energy.*

- use input from only the B diode; or
- average the inputs from the A and B diodes.

Normally, the signals from the diodes are averaged. In the case of the failure of one diode, the CARE chip can be switched to use only the input from the working diode, and the failed diode can be disconnected.

Table 10-5 summarizes an example configuration and shows how the resolution and bin size vary as functions of energy and range selection in the CARE chip. The numbers assume that the overall gain ranges are 1 (1×1), 8 (1×8), 32 (32×1), and 256 (32×8), and that a 10-bit ADC is used to digitize the data. The resolution is taken as the bin size divided by $\sqrt{12}$ with an additional factor of 2 to account for losses in clustering. The results are summarized in Figure 10-17 which shows the resolution from the digitization compared with the specified resolution for the CsI(Tl) crystals.

Since the signal is presented to the ADC chip as a DC level, the ADC chips should be able to deliver the specified number of bits of accuracy. We are exploring whether it is more cost-effective to assign a single ADC to each channel or to multiplex several channels into fewer, higher-speed ADCs.

Assuming the highest rate of sampling (4 MHz) and multiplexing, the 12 bits from each of 25 channels into a single serial output requires a capability of 1.25 Gbits/s (1.5 Gbits/s with protocol overhead). For this rate, the best technology is optical fiber.

The ADC board also performs electronic calibrations and monitors the ADC board and the preamp. A small microcontroller is placed on the board, with a low speed serial link to the outside world. This controller can read the monitoring inputs and control the data acquisition and calibration lines.

The location of the preamps and ADC boards is shown in Figure 10-18. Each group of 25 preamp cards has a cable plant of 500 wires, many of which are power and ground, and the rest of which carry signals to the ADC board serving this group of cards. A possible solution to the problem of handling this number of cables is to have a fanout board which connects

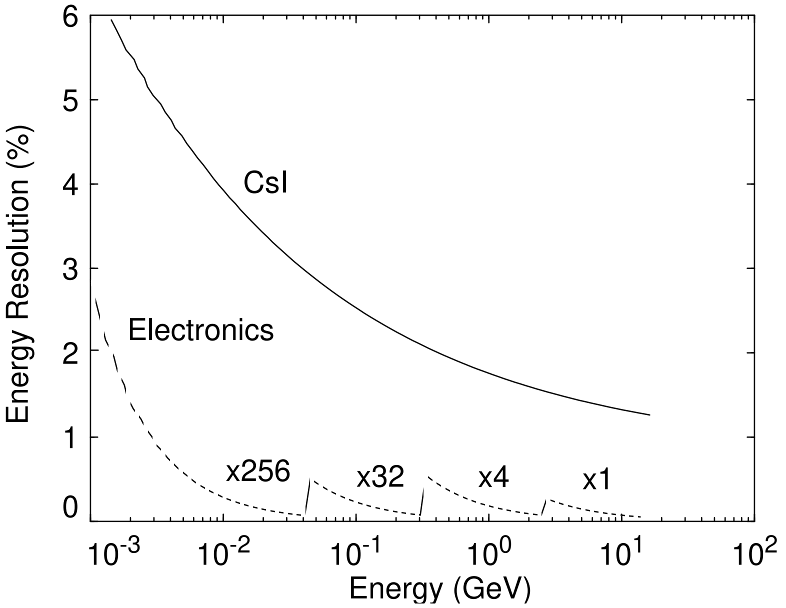

Figure 10-17. *Digitization resolution vs. CsI Resolution.*

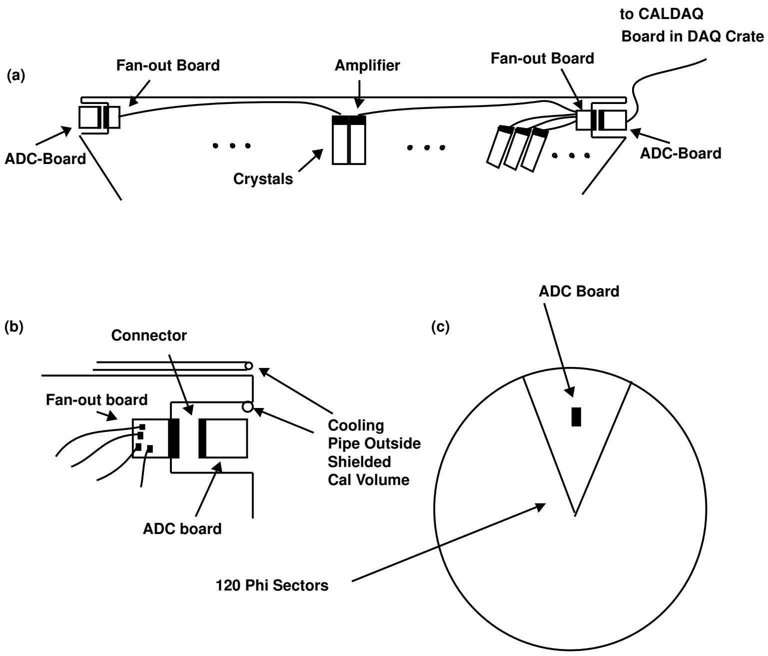

Figure 10-18. *ADC board located at the end of the barrel.*

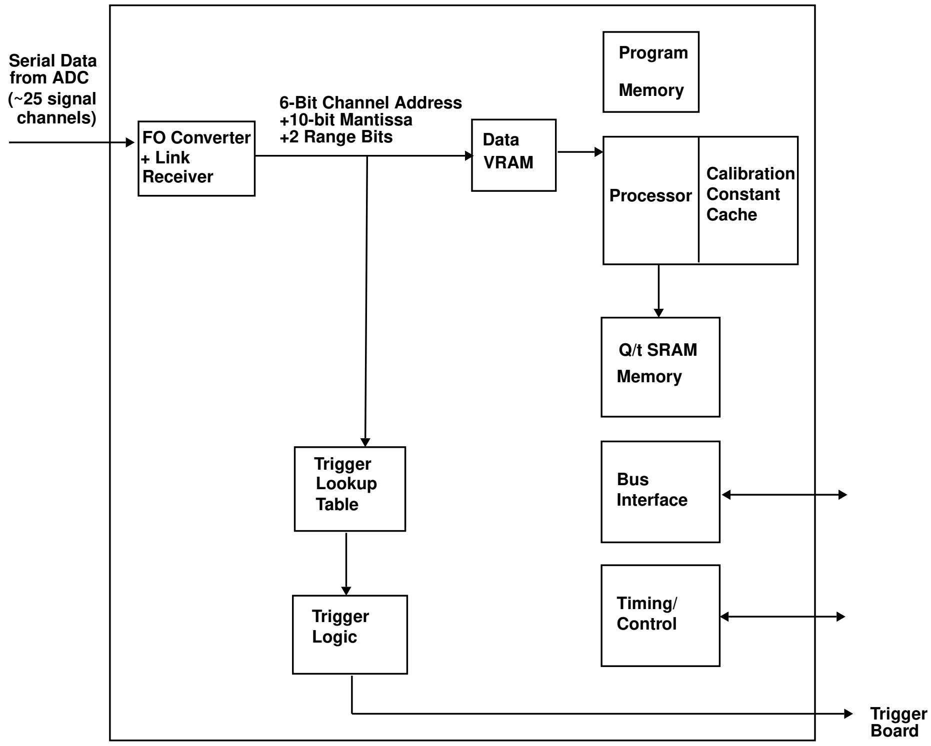

Figure 10-19. *Block diagram of readout module functions.*

the 500 cables from the preamps to a bus connector into which the ADC board then plugs. Power lines are shared across this bus, while the signals are fed to the ADC cards. This allows ADC boards to be easily replaced during shutdowns. The inputs to the ADC board from the data acquisition system are: a sampling clock, a serial output link, a slow control link, and a serial output optical fiber.

The ADC boards are placed where they can be water cooled so that the heat generated on the boards does not get transmitted to the crystals. The expected heat generation on each ADC board is 5 W, leading to a total heat load of 1.5 kW.

10.6.5 Readout Module

The readout modules for the calorimeter are located in the electronics house. They receive signals from the 280 ADC boards, provide trigger information to the Level 1 trigger, and send processed signal information to the higher level data acquisition system. This section discusses only the features unique to the calorimeter readout module. The features common to the whole data acquisition system are discussed in Section 10.10. Figure 10-19 shows a block diagram of the readout module functions. The basic functions of these modules are:

- to receive serial data from the ADC boards;
- to de-serialize the data and place it in the VRAM;
- to contain the calibration look-up table for the trigger; and
- to apply calibration corrections and to extract amplitude and time information from the waveform samples.

For de-serialization of the input data, a clock is encoded on the optical fiber signal. This guarantees that the data flowing into the VRAM are synchronized with the digitization clock, which is synchronized with the overall system clock. Some information needs to be provided to verify the channel encoding. The simplest implementation would be a single bit to flag the first channel in a stream of 25. At the cost of extra bandwidth, 5 bits could be used to send individual channel numbers with the data. Data flow into the VRAM at 1.25 Gbits/s and can be stored in the VRAM for up to 6 ms before they must be moved to more permanent memory.

For trigger purposes, the readout modules provide information about the amount of energy deposited in the 5×5 arrays of crystals which are seen by each board. This part of the module is discussed in Section 10.9.

A sophisticated algorithm to calculate the total charge and pulse height is needed to minimize the effect of the soft photon background on the resolution. Algorithms are being studied using models of the signal plus expected background. The floating-point conversion and calibration correction could be done continuously using a SRAM look-up table, but this is likely to be costly compared to doing this on demand using an onboard processor. Thus, in response to a trigger or a request containing better time information, the readout module processor executes the following steps to find the times and magnitudes of the showers: calibration constants are applied to the raw data, the charge and time of the pulse in each crystal are found, and if the magnitude is sufficiently large, this information is sent to the next level data acquisition system. Provision is also made to occasionally send the raw data to monitor the onboard algorithms.

The maximum data rate, allowing for online calculation of the time and pulse height in each crystal at a 2 kHz trigger rate, is estimated by allowing for 20 bits of charge information and 12 bits of time information. This requires 0.2 Mbytes/s to send data for all crystals on an ADC board, 3.2 Mbytes/s per readout crate, and 54 Mbytes/s for the entire calorimeter.

The design of the readout module is shared, as much as possible, with the other systems.

10.6.6 Electronic Calibration

For each channel, calibration corrections are applied on the readout module for both overall gain (Chapter 7) and the response of the analog and digitizing electronics. The latter information is provided by an electronic calibration subsystem.

In order to derive the calibration constants, the following tasks need to be performed:

- Injection of charge on the preamp: This is accomplished by sending a signal derived from an 18-bit digital-to-analog converter (DAC) on the ADC board through a capacitor onto the input of the preamp.
- Range forcing on the CARE chip: In order to understand the transition region from one range to another, the capability to force the CARE chip to remain in one gain range is essential. In this way, the operation of the automatic range selection can be verified.
- Independent control of the phase of the calibration signal with respect to the clock.
- Calibration of the CARE chip: it may be desirable to input a test signal to the CARE chip directly, and thus separate effects in the preamp from those in the CARE chip.

In addition to the calibration of the electronics, calibration with sources requires the readout of the calorimeter in an untriggered mode. This requires dedicated running using only the calorimeter and specialized software in the online processors to search for peaks in individual crystals and to create clusters. Since this calibration is performed without the beam, peak finding without a trigger should be feasible, as the design of the preamp ensures reasonably large signal/noise for source events. Both the electronic calibration runs and source calibration runs (Chapter 7) require special processing modes on the readout modules, during which the calorimeter must be in local mode.

10.6.7 Monitoring and Control

Most of the monitoring tasks are performed by the onboard microcontrollers on the ADC boards. These are interconnected with a slow low-cost link managed by an overall detector-control processor. The main variables which are monitored by the system are:

- Temperature: A tolerance of $\pm 1^\circ$C at the preamp boards and $\pm 2^\circ$C at the ADC board. A sample of preamp boards and all ADC boards are monitored.

- Bias Voltage Current: The bias voltage on and current from each photodiode are monitored. There are also provisions for measuring the total current.
- Supply Voltage: The low voltage supplies are likely to be shared across groups of ADC boards. These are monitored and safety interlocks applied to the power supplies.

Support software interprets the slow control data and provides diagnostics which allow problems to be isolated without opening the detector.

10.6.8 Research and Development

A preliminary version of the preamp IC has been constructed in a 1.2 μm BiCMOS technology. This chip uses a previous design [Bag94] modified to suit the needs of the *BaBar* calorimeter. The shaping time for this preliminary device can be modified using external components. This chip will be used in conjunction with various possible JFET candidates on the prototype calorimeter in order to evaluate the design, choice of shaping time, and performance of the preamp.

The R&D program for the ADC and readout modules centers around the following: the design and testing of the CARE chip; evaluation of candidate commercial ADCs; development of the electronic calibration system; development and testing of an appropriate VRAM control system; testing proposed trigger schemes; and benchmarking signal processing algorithms and evaluating the processor power required. A first version of the CARE chip has been built in CMOS and tested at SLAC, and a BiCMOS version of the chip capable of higher sampling rates is now in prototype fabrication. The remaining R&D issues will be the subject of parallel efforts to verify the concept and resolution obtained with the proposed methods. The new CARE chip and as many of the other components as possible will be included in the first tests of the calorimeter prototype with beam. Ongoing beam tests will study in detail the proposed production versions of all boards and modules and address system concerns which may arise only in the context of a realistic detector.

10.7 Muon System Electronics

10.7.1 Electronics Requirements

The functional requirements of the Instrumented Flux Return (IFR) electronics are to label hit strips and to associate a hit with a triggered event.

The resistive plate chamber (RPC) single counting rate is on the order of 400 Hz/ m^2 [Amb94], which is comparable to the cosmic ray rate at sea level. Assuming the probability of recording noise hits is on the order of 10^{-4}, the number of accidental hits is below 20 per event. The physics rate also is quite low, with an event generating at most a few tens of hits. Thus the IFR is a very low occupancy detector.

The readout logic will utilize the Level 1 strobe which carries a fixed latency of 10 μs with a maximum permitted jitter of $\pm 0.5\,\mu$s. The average trigger rate is less than 2 kHz, and the maximum rate the readout must tolerate is 10 kHz. The minimum time between successive triggers is 1.5 μs, and this is the time allotted to empty the front-end electronics. This small time in which to read the 42,000 channels is the main constraint on the design.

The baseline architecture presented here represents a working model which takes into account functional requirements and assumptions, and copes with constraints. The main goals of this design is to reduce the electronics cost by minimizing the number of modules and connections needed for the whole device, and to assure completely safe operation. This is accomplished by reducing the overall complexity, and by using well-defined existing technologies.

The first module (the Front-End Card, or FEC) will be located close to the detector in the iron gaps; the second module will be on the outside of the detector itself and accessible without opening the endcaps, while the last one (readout module) will be on the electronics house.

10.7.2 General Architecture

Because the RPC is practically a noiseless detector in which the rate is dominated by cosmic rays, immediate rejection of unwanted data is not as important as very rapid data collection. Buffering during the trigger latency is unnecessary because the probability of having two hits on the same strip during this time is negligible. Strip signals will be stored during the latency of the trigger decision by simply stretching them and using a fast readout card to prepare the front-end within the prescribed 1.5 μs. The IFR electronics perform as a data-driven system.

Generally, a serial readout of front-end cards is suitable when a large number of strips must be handled and the acquisition speed is not a critical parameter. However, in this case, the time allowed to empty the front-end cards is relatively short, and we plan to use a mixed serial/parallel readout design.

Conceptually, the electronics system for the IFR is a sequence of blocks performing the following functions:

- record detected signals during Level 1 strobe latency;

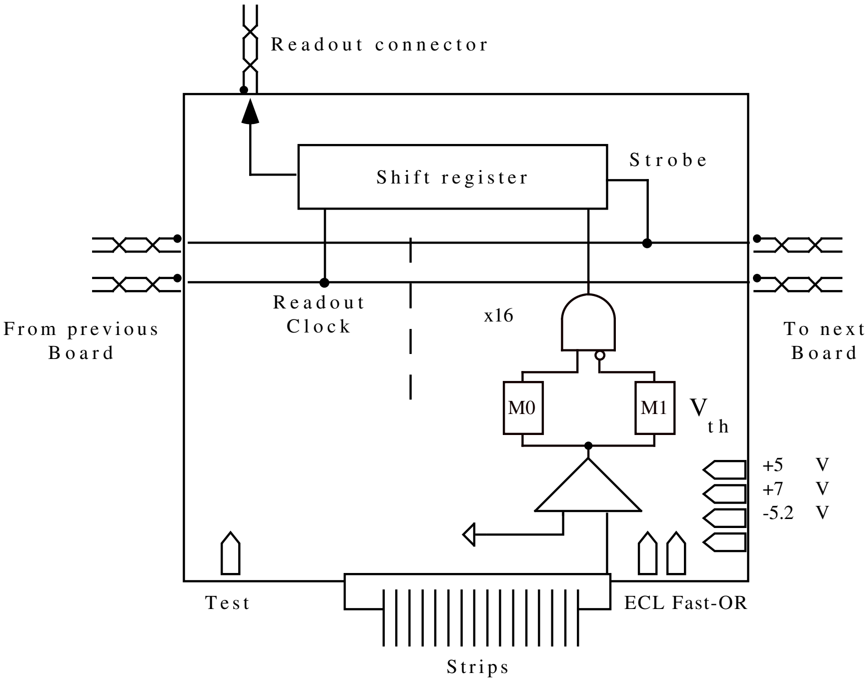

Figure 10-20. *Schematic of the Front-End Card (FEC).*

- unload the data from the front-end within the allowed time (1.5 μs), skipping empty cards;
- encode the addresses of hit strips; and
- read hit data, assemble them into a full data stream, and transfer them to a storage medium.

10.7.3 Implementation

Front-End Card (FEC)

Because of the inaccessibility of the front-end cards, they are designed with a minimum number of components. A schematic is shown in Figure 10-20. Because RPCs produce pulses consistently near 200 mV in pulse height, the discriminator stage of the FEC is not critical. A single transistor will suffice to the bring signal level to the stage of driving TTL chips [Alo94c].

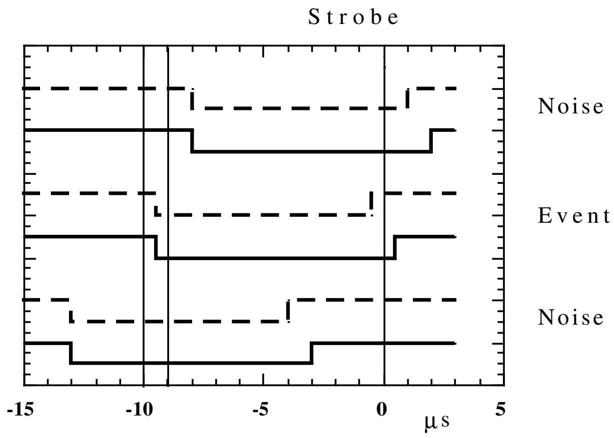

Figure 10-21. *Double one-shot logic. Only strips fired in the 1 μs window will be acquired.*

Each FEC connects directly to the ends of 16 adjacent readout strips. The output is obtained from a logical AND of two one-shots (M0 and M1), which are 9 μs and 10 μs wide, respectively. Stretching the strip signals acts both as buffering for trigger latency and for noise reduction by identifying the strip fired within the 1 μs time window (Figure 10-21). The board also provides a Fast-OR signal of the 16 strips, which can be used for time measurements, monitoring, diagnostics, and time calibration.

They are connected in a 16-card chain. Each chain is read serially as a 16-bit-wide shift register, as shown in Figure 10-22. The FECs will be installed inside the IFR gaps, very close to the RPC strips.

FIFO Buffering, Zero Suppression, and Digital Encoding (FIB)

After the Level 1 strobe is received, input data, consisting of one bit per strip, are latched into the FEC register. The chain is then read by the FIB, which is shown in Figure 10-23. This module buffers the data from the shift registers on 16 FECs and stores them in parallel in the data FIFO. The FIB module handles four different chains (1024 strips in total) and is located in a VME crate on the outside of the detector.

Each trigger has a time stamp which is read into the FIFO as a header to the data stream referring to that particular trigger. The readout chains are clocked at 12.5 MHz so that the FECs are ready for another trigger within 1.5 μs.

Independently, FIFO data are pushed into the zero suppressor and digital encoding logic, which encodes strip data into hit data. This is implemented with Xilinx LCAs and is a

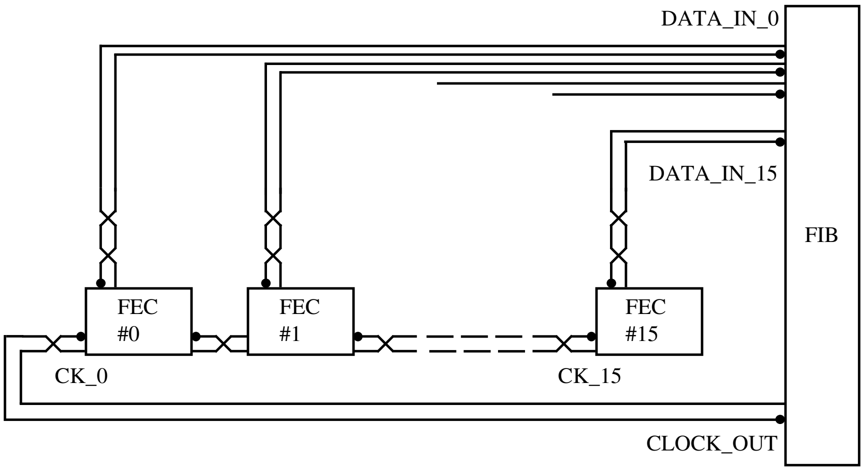

Figure 10-22. *Schematic of the parallel readout mechanism for each chain.*

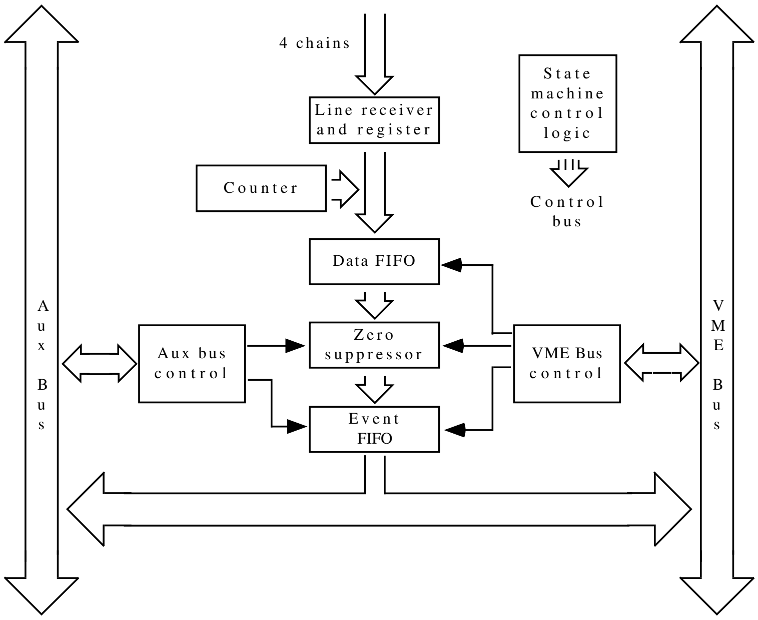

Figure 10-23. *A simplified scheme of the FIFO board.*

copy of that used in the L3 experiment at CERN [Alo94a, Alo94b, Alo94c, Alo94d, Alo94e]. Another Xilinx LCA handles VME protocol, while the state machine for the arbitration is implemented in a MACH gate array.

Readout Integration

The FIB boards are arranged in seven VME crates, three for the barrel and two for each endcap. Each crate includes a data collector module (DCM) acting as interface between the encoding module and the readout module. The DCM asks for a specified trigger number from all FIBs plugged into the same crate. It uses an auxiliary bus connection and a token-in/token-out protocol. Data flow from an FIB until no more data are available on that FIB for that trigger number. The collected data are then ready to be transferred to the readout module. This architecture is deadtimeless even under the worst background conditions.

10.7.4 Time Measurements

Time measurements can be achieved by using TDC channels with a granularity as coarse as a single layer. The measurements will be performed via common stop, VME-based, TDC modules located in the crates on the outside of the detector. The stop signal is supplied with a (synchronous) Level 1 strobe. The TDCs have a range covering a 1.5 μs time window to ensure that the entire time jitter for the trigger is contained. Provisions have to be made for the 10 μs strobe latency. When the Level 1 strobe is asserted, the TDCs are stopped and read into a FIFO buffer, associated to the right Level 1 strobe reference number, and sent to the data acquisition system. The required TDC resolution is 0.5 ns.

10.7.5 Time Calibration

Since strip signals coming from different FECs have different path lengths, their arrival time will be spread out. A time-zero measurement is thus needed in order to calibrate out different cable lengths and electronics time response.

The time calibration is performed with an external pulser system. A synchronized test pulse is injected directly in the input stage of all the FEC channels. The difference in time between the various TDC outputs provides the factor needed to calibrate the different signal delays.

10.7.6 Monitoring

The complete hardware paths can be monitored by sampling unprocessed data (spying) during data taking. Furthermore, fast-OR signals directly collected by scalars provide front-end card rates for monitoring.

10.7.7 Diagnostics

The whole electronics paths will be checked in order to identify possible failures. A test input provided on the front-end board is used to inject a test pulse on all 16 input channels. This is the same pulse signal that is used in time calibration.

10.8 Trigger Requirements and Background Rates

10.8.1 Introduction

PEP-II may give rise to rather severe backgrounds, producing up to several kilohertz of interactions with two or more charged tracks reaching the calorimeter. This rate is to be contrasted with the desired logging rate of less than 100 Hz. The trigger and data acquisition subsystems are designed to record data at no more than the latter rate under a variety of background conditions and to accommodate luminosity upgrades up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The purpose of the trigger is to reject backgrounds while selecting a wide variety of physics processes. This filtering function is performed through an increasingly sophisticated sequence of algorithms applying mostly topological cuts on increasingly refined data. Processes selected with high efficiency include final states of e^+e^- interactions ($\Upsilon(4S)$, $\Upsilon(5S)$, $c\overline{c}$, and $\tau^+\tau^-$) as well as final states of $\gamma\gamma$ and $\gamma\gamma^*$ interactions. In addition, final states such as e^+e^-, $e^+e^-\gamma$, and $\mu^+\mu^-$ are selected for use with luminosity measurements, monitoring, and calibration. Finally, for diagnostic purposes, cosmic rays, beam-induced backgrounds, and random time slices are selected at a low rate.

In addition to providing high efficiency for a large number of interesting physics processes and rejecting most background events, the trigger is designed to allow precise measurements of its efficiency. The most powerful tool for this is the use of independent selection criteria. Orthogonal selection methods are used as much as possible for those filters operating on trigger primitives, *i.e.,* on quickly derived reduced representations of the data.

Physics Data Samples	Requirements
$B\overline{B}$ from the $\Upsilon(4S)$: *CP* Channels	High Efficiency
$B\overline{B}$ from the $\Upsilon(4S)$: Non-*CP* Channels	Precise Efficiency, High Efficiency
B Physics at the $\Upsilon(5S)$	High Efficiency
Charm Physics	Precise Efficiency, High Efficiency
Continuum $q\overline{q}$	Precise Efficiency (Background Subtraction)
τ Physics: Rare Decays & Asymmetries	High Efficiency
τ Physics: Branching Ratios	Precise Efficiency
$\gamma\gamma$	Precise Efficiency

Other Data Samples	Purpose, Comments
Bhabhas: Two Detected Prongs	Luminosity, Calibration, High Efficiency
Bhabhas: One Detected Prong	Calibration, Beam Monitor
$\mu^+\mu^-$	Luminosity, Calibration
Cosmic Rays	Calibration
Beam Backgrounds	Beam Monitor, Trigger Studies
Random Events	Trigger Studies

Table 10-6. *Principal trigger requirements for each type of data sample.*

This section begins with a detailed discussion of the implications of BaBar's physics goals for the trigger. It continues with a description of trigger algorithms, and concludes with the results of simulations of the proposed trigger algorithms, including both the efficiencies for some benchmark processes and the expected trigger rates.

10.8.2 Requirements

The trigger needs to be efficient, selective, flexible, measurable, and accurate. More details are given in the *BaBar Electronics Requirements* [Req94]. The relative importance of these attributes varies for different data sets. For asymmetry measurements and searches, it is more important to have a high efficiency than a well-known efficiency. For branching ratio and cross-section measurements, precisely measured efficiencies are crucial. The principal trigger requirements for each data set are listed in Table 10-6. In addition to the key physics events, a number of other events are recorded for uses such as luminosity determination, detector calibration, beam monitoring, and trigger performance studies. These impose additional requirements on the trigger and are listed at the bottom of the table.

Because of their relatively large multiplicities and visible energies, it is straightforward to achieve high trigger efficiencies for Υ and $c\overline{c}$ events. The challenge for the trigger designer lies in detecting low multiplicity $\tau^+\tau^-$ and $\gamma\gamma$ events while keeping the rate of accepted background events within the data acquisition limit, so as not to incur deadtime. To ensure high efficiencies in offline analysis and large background rejection power, trigger algorithms need to have sharp efficiency turn-on curves. For example, the efficiency as a function of cluster energy should approximate a unit step function as closely as possible to allow a highly selective trigger.

Because PEP-II machine conditions will vary over the course of the experiment, the trigger needs to be flexible. Under temporarily severe background conditions, saturation of the bandwidth should result in deadtime in such a way that a diminished number of whole events are collected rather than a larger set of partial events. When backgrounds are severe but data are useful for physics, prescaling factors can be applied to selection streams to ensure that the available bandwidth is filled with interesting events. This can be done on a relatively short timescale. It will be straightforward to change the global logic combinations of the selection streams in each of the trigger levels, and it will also be possible to vary selection parameters. Actual change of trigger selection algorithms is foreseen to occur rarely.

There are several ways to facilitate precise offline measurements of the trigger efficiency. First, the geometry of the detector and the readout electronics for each system contributing to the trigger are designed with trigger considerations in mind. Second, the trigger and data acquisition system is designed to:

- use orthogonal selection criteria;
- record all trigger primitives and decisions for accepted events;
- accept prescaled numbers of events with looser criteria;
- accept a prescaled number of rejected events;
- have a sharp efficiency turn-on as a function of the selection variables; and
- be simple to model with simulations.

Finally, the online system will be designed so that the recorded trigger primitives and decisions are readily available for offline analysis.

Use of the trigger results for precision physics measurements depends crucially on the trigger system working as a whole. Like the other systems that deliver data to the data acquisition, the trigger system should be designed with sufficient monitoring, calibration, and diagnostic testing capabilities so that trigger decisions and trigger data are known to be correct.

10.8.3 Backgrounds and Trigger Rates

The trigger architecture and selection algorithms designed to meet the requirements are based on simulation studies of the physics, backgrounds, and the *BaBar* detector.

The total trigger output rate is the sum of the rates for the physics and diagnostic processes listed in Table 10-6 and the rates for the background processes described in Chapter 12.

The main background processes that produce charged tracks traversing the drift chamber and calorimeter are:

- physics backgrounds from asymmetric Bhabhas with only one detected charged track and two-photon events below $\gamma\gamma$ energies of interest;
- beam-induced backgrounds from hadronic interactions of lost beam particles, electromagnetic interactions of lost beam particles, and beam-gas collisions at the interaction point; and
- cosmic rays.

The main sources of occupancy in the vertex detector, drift chamber, and calorimeter are:

- particles from small-angle radiative Bhabhas, and
- beam-induced background from lost beam particles.

The rate of charged particles and the occupancy of physics backgrounds scale with the luminosity, whereas those of beam-induced backgrounds scale with the beam current and the pressure of the residual gas inside the PEP-II beam pipe.

Physics backgrounds are evaluated at the design luminosity, and the trigger and data acquisition system must be upgradable for higher luminosities. Because of their variable nature and because of uncertainties in the predictions, beam-induced backgrounds pose more stringent design requirements. The trigger and data acquisition system is required not to produce substantial impact on the physics program as long as data are useful. Based on simulation of the occupancies in the vertex detector as functions of lost-particle rate, pattern recognition becomes seriously compromised at ten times the nominal background rate. The simulation is described in Chapter 12.

10.8.4 Definitions of Trigger Filter Concepts

Event Selection

Event signatures are used to separate signal from background. Combinations of the following global event properties are used in the trigger filters:

- charged track multiplicity;
- calorimeter cluster multiplicity;
- event topology; and
- primary vertex position.

These selection criteria have associated thresholds for the following parameters: charged-track transverse momentum (p_t), energy of calorimeter clusters (E_{clus}), solid angle separation (ϕ and θ), track match quality (for electrons and muons) and primary vertex impact parameter (ρ_1 and z_1, respectively). The trigger definition can contain selection criteria that differ only by the values of thresholds, such as the p_t or E_{clus} threshold values. A small fraction of random beam crossings and events with failed triggers are selected for diagnostic purposes.

Selection Filters

Simulation studies both of processes of interest in the areas of B, τ, and $\gamma\gamma$ physics [Kra94b, Kra94a, Lev94a, Lev94b, Lev94c, Lev94d, Sny94], and of background events in the BaBar detector, have been used to guide the choice of trigger algorithms.

Table 10-7 lists "trigger objects" which are potentially definable at Level 1 and which have been determined to be useful in constructing global trigger definitions. Vertex detector information is available at higher levels. For a given trigger level, the global selection filter is a logical OR of a number of specific selection filters, each of which is the result of a boolean operation involving any of the ingredients available at that level. The types of selection filters may be arranged in the following hierarchy:

1. Multiplicity of drift chamber charged particles with implicit p_t and $\cos\theta$ cuts and explicit ϕ-separation cuts;
2. Multiplicity of calorimeter clusters with explicit E_{clus} and implicit $\cos\theta$ cuts, and explicit θ- and ϕ-separation cuts;

Tracks in the Drift Chamber	
A tracks	Tracks that reach the outer layers of the drift chamber. For $B = 1.5\,\mathrm{T}$, this implicitly leads to approximately $p_{\mathrm{t}} > 0.210\,\mathrm{GeV}/c$. These tracks are contained in the solid angle of approximately $-0.81 < \cos\theta < 0.90$.
A′ tracks	A tracks that satisfy an explicit cut on the reconstructed p_{t}.
B tracks	Tracks that reach the middle layers of the drift chamber. This implicitly leads to approximately $p_{\mathrm{t}} > 0.130\,\mathrm{GeV}/c$. The solid angle acceptance for these tracks that can exit the middle of the drift chamber end cone is substantially larger than for A tracks, $-0.92 < \cos\theta < 0.96$.
Clusters in the Calorimeter	
M clusters	Clusters that have reconstructed energy deposits consistent with minimum ionizing particles (MIP), or larger, in the calorimeter barrel or endcap. The required value is near 150 MeV. A minimum ionizing μ^- deposits 181^{+20}_{-8} MeV, averaged over the solid angle of the barrel calorimeter, where the limits indicate the variation in energy deposit due to trajectories and varying crystal lengths [Kra94a]. For such a muon to reach the calorimeter barrel, it is implicitly required to have approximately $p_{\mathrm{t}} > 0.250\,\mathrm{GeV}/c$. The solid angle acceptance is approximately $-0.78 < \cos\theta < 0.96$. The cluster energy threshold is adjustable.
G clusters	Clusters with energies above those of M clusters; *i.e.,* $E_{\mathrm{clus}} > 0.200\,\mathrm{GeV}$. The cluster energy threshold is adjustable.
E clusters	Clusters with energies large enough to be produced by $e^+e^-\gamma$ or $\gamma\gamma^*$, *i.e.,* approximately $E_{\mathrm{clus}} > 1.0\,\mathrm{GeV}$. The cluster energy threshold is adjustable.
Matches Between the Drift Chamber and Calorimeter	
A-M matches	A tracks correlated in ϕ with barrel M clusters.
A′-M matches	A′ tracks correlated in ϕ with barrel M clusters.
B-M matches	B tracks correlated in ϕ with barrel or endcap M clusters.

Table 10-7. *Trigger objects.*

3. Special triggers for calibration, monitoring, and diagnostics;
4. Multiplicity of charged particles correlated between the drift chamber and calorimeter M clusters, with explicit matching in ϕ;
5. Multiplicity of drift chamber charged particles with implicit $\cos\theta$ cuts, and explicit p_{t} and ϕ-separation cuts; and
6. Multiplicity of charged particles correlated between the vertex detector and drift chamber with explicit matching in ϕ or explicit cuts on ρ and z.

The list continues with more refined selection cuts applied to the full data set available in Level 3.

At each hierarchical step, there may be a number of selection filters (trigger lines), differing in whether certain cuts are applied or in the values of parameters used to make cuts. Let the isolated symbol for an object represent the number of such objects, *e.g.,* A is the number of A tracks. Assume that counting and geometry cuts are such that no physical object will fill a double role in the trigger. Then, using the notation of Table 10-7, category (2) might include: $\mathrm{M} \geq 2$ with minimal separation cuts; one or more versions of $\mathrm{G} \geq 1 \cdot \mathrm{M} \geq 1$ with minimal separation cuts; or $\mathrm{G} \geq 1 \cdot \mathrm{M} \geq 1$ with appreciable ϕ separation. Each of the trigger lines will be subject to individual prescaling, which may have to be used to reduce rates for the more open definitions, while collecting all events satisfying more stringent definitions. The global selection OR at each level is formed after any such prescaling. Some specific filters will be evaluated for their effects on signal and background events in the following subsection.

10.8.5 Performance of Some Simple Triggers

Benchmark Physics Processes

In order to evaluate the efficiency of selection algorithms, a set of physics processes has been identified to serve as trigger benchmarks. As explained in Chapter 3, the *BaBar* physics program focuses on studies using $\Upsilon(4S)$, $\Upsilon(5S)$, $q\overline{q}$, $\tau^+\tau^-$, $\gamma\gamma$, and $\gamma\gamma^*$ events. While the primary mission of the *BaBar* experiment is the study of *CP* violation in B decays, the trigger is designed to accept a large number of processes that may be important. The trigger benchmarks represent processes that are difficult to distinguish from background. Other similar processes have, or are expected to have, higher efficiencies. Here we use Monte Carlo information to study the effectiveness of some simple triggers on these benchmarks.

For $B^0\overline{B}^0$ *CP* studies, the most difficult tag to trigger on is $\overline{B}^0 \rightarrow \mu\mathrm{X}$ because of its low charged-track multiplicity and low cluster energy compared to the electron or kaon tags.

The low-multiplicity decay $B^0 \to \pi^+\pi^-$ was chosen as a *CP* study decay mode to test the efficiency of our simple particle counting trigger algorithms. The high-multiplicity decay $B^0 \to D^+D^-$ was chosen to test algorithms that require fast tracks.

For $\tau^+\tau^-$ studies, events in which both taus decay to one charged track are the most difficult to select. To measure the muon branching ratio, events with $\tau^- \to \mu^-\bar{\nu}\nu$ accompanied by $\tau^+ \to e^+\nu\bar{\nu}$ are used. This is a worst-case decay for the purpose of the benchmark table, as the measurement is not expected to be statistics limited. The other $\tau^+\tau^-$ benchmark is an untagged event for which each tau decays to one charged pion, used to measure spin correlations for V−A tests.

The $\gamma\gamma$ benchmarks are straightforward: two-pion events that are either both charged or both neutral. For two-charged-hadron final states, one can normalize the hadronic cross section to the better known muon-pair cross section. The $\gamma\gamma^*$ benchmarks with two kaons and one pion in the final state, produced via an unspecified 1^{++} resonance, represent searches for exotic mesons.

Benchmark Efficiencies and Background Rates

An inclusive global trigger definition that efficiently selects each of the processes in Table 10-6 while rejecting some backgrounds requires a minimum of two charged tracks in the drift chamber, two M clusters, or one E cluster in the backward direction [Bol91]. The orthogonality of the two first criteria makes it straightforward to measure the efficiency of the filter.

The following discussion focuses on defining a flexible open trigger with high efficiency. Based on the first two criteria listed below, it is able to handle a large range of potential beam-induced backgrounds.

There are three versions of the two-track trigger to consider: $\mathrm{A} \geq 2$, $\mathrm{A} \geq 1 \cdot \mathrm{B} \geq 1$, and $\mathrm{B} \geq 2$. The preferred definition is $\mathrm{A} \geq 1 \cdot \mathrm{B} \geq 1$ (the so-called one and one-half track trigger) because it has larger solid angle acceptance than $\mathrm{A} \geq 2$ and accepts fewer looping tracks than $\mathrm{B} \geq 2$. Simulations show that $\mathrm{A} \geq 2$ is only ∼80% as efficient as $\mathrm{A} \geq 1 \cdot \mathrm{B} \geq 1$ for inclusive tau decays [Bol91, Kra94b]. One soft charged particle can loop back, resulting in two detected B tracks. This problem is less severe with the $\mathrm{A} \geq 1 \cdot \mathrm{B} \geq 1$ definition.

Efficiencies and output rates for the two-particle trigger were simulated for two global-trigger input lines, D_2 and C_2. The default two-track trigger D_2 uses the $\mathrm{A} \geq 1 \cdot \mathrm{B} \geq 1$ definition and is simulated. The default two-cluster trigger C_2 requires $\mathrm{M} \geq 2$, both with $E_{\mathrm{clus}} > 0.150\,\mathrm{GeV}$, where the chosen value for the cluster energy threshold is a compromise between efficiency and background rate. To suppress pile-up of soft photons due to electromagnetic interactions

Process	D_2	C_2	$C_2 + D_2$
$B^0 \to \pi^+\pi^-$, $\overline{B}^0 \to \mu X$	1.00	1.00	1.00
$B^0 \to D^+D^-$, $\overline{B}^0 \to X$	1.00	1.00	1.00
$\tau^+ \to e^+\nu\bar{\nu}$, $\tau^- \to \mu^-\bar{\nu}\nu$	0.72	0.73	0.79
$\tau^+ \to \pi^+\bar{\nu}$, $\tau^- \to \pi^-\nu$	0.85	0.74	0.89
$\gamma\gamma \to \pi^+\pi^-$, $W_{\gamma\gamma} > 2\,\mathrm{GeV}/c^2$	0.63	0.63	0.63
$\gamma\gamma \to \pi^0\pi^0$, $W_{\gamma\gamma} > 2\,\mathrm{GeV}/c^2$		0.86	0.86
$\gamma\gamma^* \to X(1420) \to KK\pi$	0.48	0.62	0.62
$\gamma\gamma^* \to \eta(550) \to \pi^+\pi^-\pi^0$	0.07	0.21	0.21

Table 10-8. *Open trigger efficiencies for benchmark physics processes. The efficiency denominator includes* all *events generated over the full* 4π *solid angle. The simulation used 1000 events.* $W_{\gamma\gamma}$ *is the mass of the* $\gamma\gamma$ *system.*

of lost beam particles, each calorimeter channel is required to have an energy deposit above a threshold of 10 MeV. An open global trigger is defined as $C_2 + D_2$.

Calculated efficiencies for benchmark physics processes using the defined trigger lines are listed in Table 10-8. These efficiencies are found to be very high for each of the benchmark processes. Estimated open-trigger output rates due to the processes listed in Table 10-6 and in Section 10.8.3 are shown in Table 10-9. The simulation is described in Section 10.9.5.

Using the open trigger, the maximum event-building rate goal is not satisfied at $\times 10$. Thus, the open trigger would need to be prescaled and a tighter criterion would need be applied when the total output rate exceeds 2.0 kHz. The best solution to this problem—*i.e.,* with the least impact on two-prong physics processes—is to require an A$'$ track instead of an A track and a G cluster instead of one of the M clusters. These are the D_2' and C_2' trigger lines. To suppress lost-particle electromagnetic backgrounds, the C_2' line further required the G cluster to be back-to-back with an M cluster. The efficiency and rate for the two trigger lines are given in Tables 10-10 and 10-11, respectively.

10.9 Level 1 Trigger

10.9.1 Overview

The first level trigger architecture provides the Level 1 strobe for the front-end buffers and reduces the event output rate for feature extraction and event building. The trigger achieves

Process	D_2	C_2	$C_2 + D_2$
Physics:			
$\Upsilon(4S)$	12 Hz	12 Hz	12 Hz
$q\overline{q}$	8 Hz	8 Hz	8 Hz
$\mu^+\mu^-$	2 Hz	2 Hz	2 Hz
$\tau^+\tau^-$	2 Hz	2 Hz	2 Hz
Bhabha	100 Hz	100 Hz	100 Hz
$\gamma\gamma$ signal	3 Hz	3 Hz	3 Hz
$\gamma\gamma$ background	50 Hz	50 Hz	50 Hz
Backgrounds:			
Hadronic nominal	300 Hz	140 Hz	380 Hz
Electromag. nominal	< 60 Hz	230 Hz	230 Hz
Beam-gas 1 nTorr	< 4 Hz	< 2 Hz	< 6 Hz
Hadronic $\times 10$	3000 Hz	1400 Hz	3800 Hz
Electromag. $\times 10$	450 Hz	3800 Hz	4000 Hz
Beam-gas 10 nTorr	< 40 Hz	< 20 Hz	< 60 Hz
Cosmic rays	80 Hz	410 Hz	440 Hz
Total nominal	620 Hz	970 Hz	1200 Hz
Total $\times 10$	3700 Hz	5800 Hz	8500 Hz

Table 10-9. *Open trigger rates. The background rates are subject to large uncertainties. The hadronic and electromagnetic backgrounds are due to lost beam particles. The rates of the first group of processes are dependent on the luminosity. The rates of the second and third groups are dependent on the product of the beam current and the residual pressure in the beam pipe. The cosmic rate is fixed.*

these goals through the use of a fast hardware trigger operating on primitives, provided via point-to-point links by the drift chamber and calorimeter. The drift chamber and calorimeter triggers feed information to the global trigger, which issues the Level 1 strobe.

The structure of the Level 1 trigger is shown in Figure 10-24. Data from the drift chamber and calorimeter are converted into trigger primitives by local logic on the readout modules. There are three steps of trigger logic between the formation of primitives and the global Level 1 decision that produces the Level 1 accept signal: particle finders that identify drift chamber tracks and calorimeter clusters, particle counters that identify the number of distinct particles, and a track matcher that associates drift chamber tracks with calorimeter clusters.

Process	D_2'	C_2'	$\mathrm{C}_2' + \mathrm{D}_2'$
$B^0 \to \pi^+\pi^-$, $\overline{B}^0 \to \mu\mathrm{X}$	0.99	0.75	0.99
$B^0 \to D^+D^-$, $\overline{B}^0 \to \mathrm{X}$	0.97	0.80	0.99
$\tau^+ \to e^+\nu\bar{\nu}$, $\tau^- \to \mu^-\bar{\nu}\nu$	0.65	0.43	0.68
$\tau^+ \to \pi^+\bar{\nu}$, $\tau^- \to \pi^-\nu$	0.78	0.36	0.80
$\gamma\gamma \to \pi^+\pi^-$, $\mathrm{W}_{\gamma\gamma} > 2\,\mathrm{GeV}/c^2$	0.62	0.51	0.63
$\gamma\gamma \to \pi^0\pi^0$, $\mathrm{W}_{\gamma\gamma} > 2\,\mathrm{GeV}/c^2$		0.76	0.76
$\gamma\gamma^* \to X(1420) \to KK\pi$	0.18	0.29	0.31
$\gamma\gamma^* \to \eta(550) \to \pi^+\pi^-\pi^0$	0.05	0.06	0.06

Table 10-10. *Safe trigger efficiencies for benchmark physics processes. The efficiency denominator includes all events generated over the full 4π solid angle. The thresholds were $p_t^{\mathrm{A}} > 0.600$ GeV/c for D_2' and $E^1_{\mathrm{clus}} > 0.600$ GeV for C_2'. Clusters were required to be separated by at least 150° in ϕ. $\mathrm{W}_{\gamma\gamma}$ is the mass of the $\gamma\gamma$ system.*

Process	D_2'	C_2'	$\mathrm{C}_2' + \mathrm{D}_2'$
Physics Signal	47 Hz	124 Hz	127 Hz
$\gamma\gamma$ Background	50 Hz	20 Hz	50 Hz
Hadronic nominal	90 Hz	10 Hz	100 Hz
Electromag. nominal	< 60 Hz	< 60 Hz	< 60 Hz
Beam-gas 1 nTorr	0 Hz	< 1 Hz	< 1 Hz
Hadronic ×10	900 Hz	120 Hz	980 Hz
Electromag. ×10	< 220 Hz	220 Hz	220 Hz
Beam-gas 10 nTorr	< 2 Hz	< 10 Hz	< 12 Hz
Cosmic Rays	50 Hz	2 Hz	50 Hz
Total nominal	300 Hz	220 Hz	390 Hz
Total ×10	1300 Hz	500 Hz	1400 Hz

Table 10-11. *Safe trigger rates. The background rates are subject to large uncertainties. The D_2' line required $p_t^{\mathrm{A}} > 0.600$ GeV/c and the C_2' line required $E^1_{\mathrm{clus}} > 0.600$ GeV. Clusters were required to be separated by at least 150° in ϕ. The total output rate is the sum of the safe trigger rate, the prescaled open trigger rate, 10 Hz of prescaled asymmetric Bhabha triggers, and 5% of special triggers.*

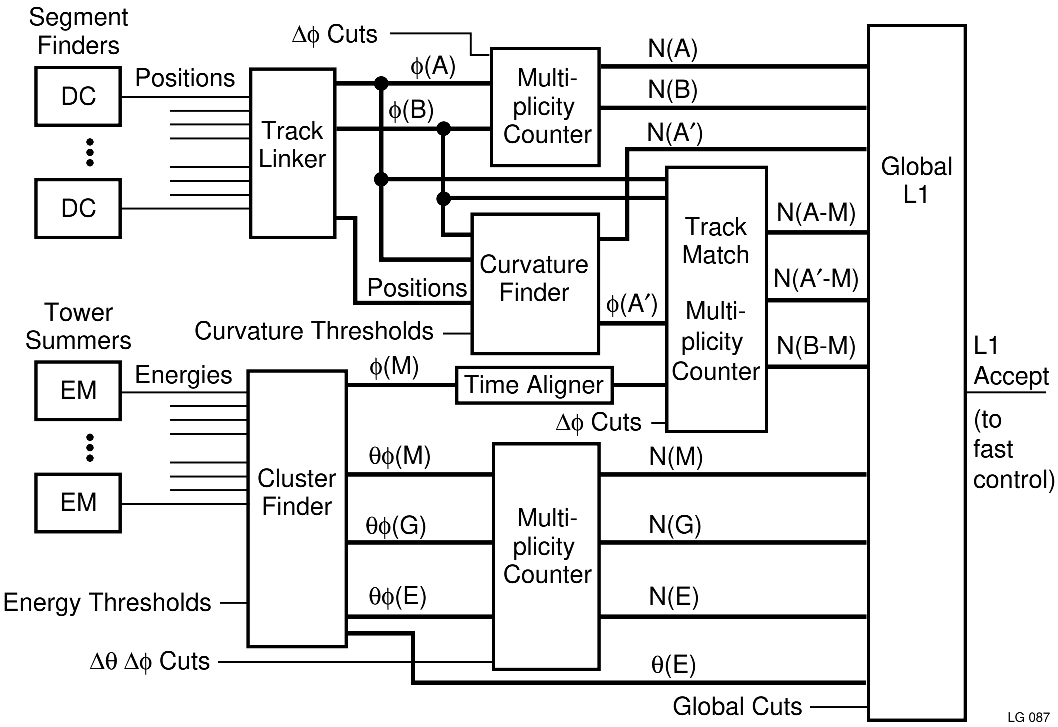

Figure 10-24. *Level 1 trigger structure, data flow, and decision logic.*

The different timing structure of the drift chamber and calorimeter primitives and the delay associated with the inputs to the track matcher require one stage of time alignment.

A number of variable selection cuts are available to these trigger logic steps. They could be tuned during initial running but should be stable during subsequent running so as not to complicate offline analysis. The global decision parameters are also flexible in the same way, to accommodate changing data-taking conditions. During more rapidly changing beam conditions, the Level 1 accept rate is controlled by prescaling certain components of the selected events. Information from the Level 1 global trigger is used for detailed accounting of prescaling factors and deadtimes.

In addition, there are important interfaces of the Level 1 trigger to the data acquisition, fast control, and online systems. These interfaces are discussed after a detailed description of the trigger decision implementation.

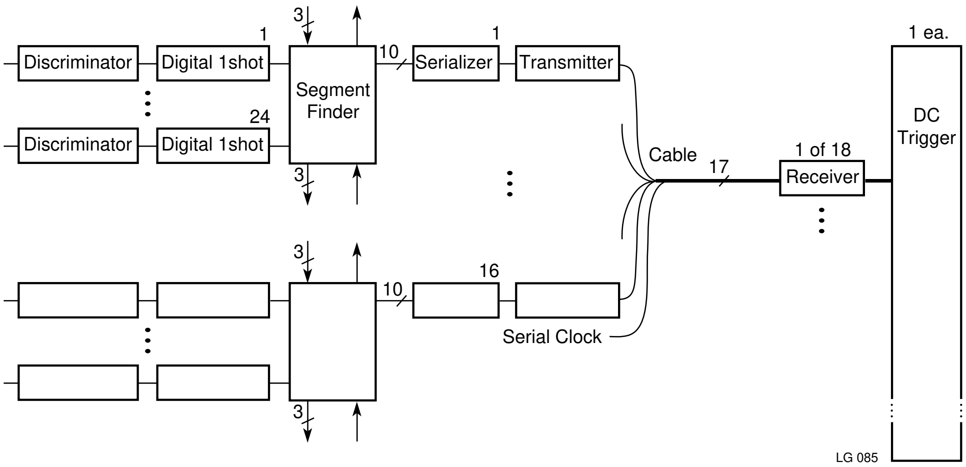

Figure 10-25. *Level 1 drift chamber trigger electronics for primitives.*

10.9.2 Drift Chamber Trigger

The drift chamber trigger performs three functions: local primitive formation, transmission and gathering of primitives, and global track finding. The structure of the drift chamber trigger is shown in Figure 10-24 with details relating to the primitives shown in Figure 10-25. The trigger algorithms consist of two track finders. The first identifies local track segments using information from readout cards that receive signals from 20 neighboring wire-chamber cells in the same superlayer. The second links these segments together to form track candidates. This section describes the algorithms and a preliminary electronics design for the baseline axial-stereo drift chamber [Kra94b].

Primitive Finders: Local Segment Finding

The segment finder algorithm is closely tied to the electronics implementation [Bol91, SLA93]. As is shown in Figure 10-7, trigger primitives are formed on the readout modules, which receive analog signals from 24 channels comprising a 6×4 supercell. To accommodate the maximum drift time, these signals are discriminated and converted to 600 ns-long signals by digital one-shots. Such signals from the 24 channels on the card are input, together with four signals from the neighboring readout modules, to a field-programmable gate array that finds track segments at a rate of one decision per 250 ns. Note that drift time information is not used.

Possible track-segment patterns consist of a hit in each of the four layers, with each hit being an immediate neighbor of the hit in the layer inside. An example of such a pattern is shown in

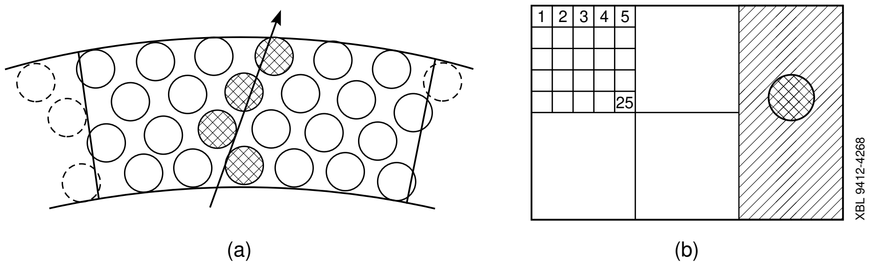

Figure 10-26. *Level 1 trigger primitives for the (a) drift chamber and (b) calorimeter. A drift chamber track produces patterns like the one shown and uses majority logic to allow three of four layers hit. For the calorimeter, energies from 25 blocks are summed locally to form towers. The cluster sums shown for the example energy deposit are formed from the primitives in a central location.*

Figure 10-26(a). The elimination of sideways tracks provides background rejection without significant loss in physics efficiency. Majority logic is implemented to guard the algorithm against inefficiencies. Patterns containing three of the four hits required for a pattern also produce valid segments. The segment finder generates 12 output bits, representing ϕ position in the 6×4 supercell. The local angle of the segment is not reported.

Transmission Links: Segment Gathering

Transmission of drift chamber primitives is via a dedicated point-to-point serial link, as outlined in one implementation of the calorimeter trigger [Wun94]. Using a serial link reduces the number of wires needed and makes it possible to gather the signals in one location.

Particle Finders: Global Track Linking

A simple and efficient track linking algorithm is the Binary Link Track finder (BLT), developed for the CLEO II trigger [CLE91] by K. Kinoshita [Kin89]. The method is to start at the second superlayer and move radially outward. At each superlayer, a segment is scored only if it is hit and any of its three immediate neighbors in the previous superlayer have been hit. Here, hit refers to a segment-finder hit. At the last superlayer, denoted a counting layer, any surviving scored segments represent tracks that traveled continuously from the first superlayer to the counting layer. The requirement of three neighbors imposes loose curvature cuts on tracks, while being forgiving to superlayer stereo stagger effects.

For *BaBar*, majority logic is incorporated by allowing combinations of four out of five superlayers to contribute segments. The B-track linker uses all five permutations of four or five of the the inner five superlayers, while the A-track linker uses all 36 permutations of four or five of the inner superlayers together with four or five of the outer superlayers.

The BLT algorithm was chosen because it prevents double counting and because it avoids searching for tracks in "roads." In the current drift chamber geometry, the latter method would require a sector with $\phi = 74°$ (one-fifth of 2π) in order to just barely contain a track with $p_{\mathrm{t}} = 180\,\mathrm{MeV}/c$. To contain any track, one would need to search in twice that region and then account for overlaps. Furthermore, unlike most road-based algorithms, the BLT is insensitive to the sign of the electric charge of the track. Because it can be implemented in combinational logic hardware, the BLT is also fast and inexpensive.

Particle Counters: Global Track Multiplicity Finding

The BLT outputs are used to count distinct tracks using simple multiplicity logic. The output of the BLT is two digital strings with ϕ positions of segments in the fifth and tenth superlayers, representing tracks passing through these points. The logic checks to see if there are at least two B tracks separated by an adjustable number (taken to be two) of supercells in the fifth superlayer (for a ϕ isolation of at least 22°). For events with two or more A tracks, ϕ isolation is checked in superlayer ten. The multiplicities of distinct A tracks and B tracks are reported as 3-bit numbers to the global trigger.

Transverse Momentum Determination: Curvature Finding

Using ϕ positions of the axial layers of the drift chamber, it is straightforward to design a Level 1 drift chamber trigger that reports multiplicities above crude p_{t} thresholds [CLE91]. The thresholds can be used to suppress the most severe trigger backgrounds from lost-beam-particle interactions, as shown in Section 10.8.

Drift Chamber Trigger for the All-Stereo Option

An all-stereo chamber with fully interleaved stereo layers is also being considered. This would require elaboration of the tracking trigger. Logical superlayers would have to be formed from alternating layers. A possible choice is six three-layer superlayers in each stereo view. The definition of logical supercells then leads to more than 300 unique segment-finding algorithms with increased interconnectivity of the readout modules.

The apparent curvature of tracks in one view varies with the track's dip angle. This would be largely corrected by summing the curvatures of tracks found in each view. For events with small multiplicities, all pairs would be summed in parallel. Events with large multiplicities would be accepted without the p_{t} requirement.

10.9.3 Calorimeter Trigger

A calorimeter trigger based on towers (groups of contiguous crystals) will be more efficient than a trigger which uses only the energy deposits in individual crystals [Kra94a] because the center crystal of a shower will contain at most about 80% of the shower energy [SLA93]. Besides the improvement in efficiency, the better estimate of the total energy deposited in a tower compared with a crystal means that any energy threshold cut applied in the trigger is more precise, which is not only better for background rejection, but is also important for understanding and modeling the trigger response accurately.

The algorithm for the calorimeter Level 1 trigger [Kra94a] is summarized in this section. It is based on fixed towers of mostly 5×5 crystals and performs three functions: local tower energy summing to form primitives, transmission of the primitives and cluster reconstruction, and central particle finding and counting. The formation of primitives occurs on the calorimeter readout modules (Section 10.6). The primitives are then transmitted to a central trigger crate where cluster reconstruction is carried out. Finally, the results of the clustering are compared and physical particles are found.

Primitive Finders: Local Tower Summing

Since the Level 1 calorimeter trigger is based on towers, the most straightforward arrangement is to organize the electronics by tower at the board level. Each tower, consisting of the signals of a 5×5 group of crystals [Bol91, Kra94a], will be read out by a single calorimeter readout module (Section 10.6), and the calorimeter trigger primitives will be constructed from tower sums formed on the board. With the numbers of crystals listed in Chapter 7 and a 5×5 tower structure, there will be 24 towers in ϕ and ten towers in θ along the barrel, with the towers at the backward edge of the barrel containing only 4×5 crystals. For the endcap, there is a 20-fold ϕ symmetry in the three superlayers, with each ϕ sector containing 45 crystals. Two towers will be formed from each ϕ sector, with five unused channels shared between each pair of boards. This gives a total of $24\times10+20\times2 = 280$ towers. The trigger sampling rate is fixed to be the same as the data digitization rate (4 MHz).

For each time sample, the 25 channels of data are calibrated and converted to a cruder energy scale of 8 or 16 bits using a look-up table in SRAM. This table allows a large degree of flexibility; an individual channel threshold cut can be trivially applied to remove low-energy backgrounds. The 25 outputs are summed using a programmable logic device (PLD) to give a total tower energy for each sample, again in 8 or 16 bits, and the sums are transmitted to a central trigger crate. These sums form the calorimeter trigger primitives, although it is possible that, to obtain a sufficiently accurate time and energy for the tower, further processing might be performed on the readout modules rather than in the trigger crate. However, assuming the raw sums are sent, the data rate is 4 MHz$\times$ 8–16 bits or 4–8 Mbytes/s.

Transmission Links: Tower Gathering

The transmission of the calorimeter primitives is done via a dedicated point-to-point serial link. The 4–8 Mbytes/s rate is reasonably low and is straightforward for a TAXI or similar link [Wun94]. The primitives are collected together either on several cards in one trigger crate, or possibly all on a single card.

Particle Finders: Global Cluster Formation

Due to energy sharing between towers, the towers are considered in neighboring pairs, called dominos [Ell92], as indicated in Figure 10-26(b). All possible dominos are formed, and the total energy of each domino cluster is used for the trigger. There will be three separate energy thresholds for clusters in the trigger: an M cluster (above a low energy consistent with a minimum ionizing particle, *e.g.,* 150 MeV), a G cluster (higher than a MIP, *e.g.,* 600 MeV), and an E cluster (consistent with a Bhabha electron). These correspond to the trigger objects in Table 10-7. The multiplicities of each object required for a trigger are programmable in the central trigger logic. Due to the differing crystal lengths in different regions of the calorimeter (Table 7-3), the MIP signal size will depend on θ, and different energy thresholds can be used in different towers if necessary.

Particle Counters: Global Cluster Multiplicity Finding

To make isolation requirements on the clusters, the cluster pairs are grouped together in superclusters of between 9 and 12 towers each. For example, with eight superclusters in ϕ and three in θ in the barrel, and eight in ϕ and one in θ in the endcap, there will be a total of 32 superclusters. By counting the number of clusters above threshold in each supercluster, and comparing which superclusters have been hit, simple isolation requirements such as two non-neighboring superclusters can be easily set.

One subtlety is that low-angle Bhabhas with both the electron and positron hitting the calorimeter have a cross section on the order of 30 nb [Eis90], giving a rate of $\sim$100 Hz at design luminosity. These are prescaled.

10.9.4 Global Trigger

The global Level 1 trigger consists of two pieces: a track matcher and a global decider.

Track Matcher

The track matcher uses ϕ maps from the drift chamber and calorimeter triggers to find the match multiplicity. There are two types of matches: a track to barrel M cluster, and B track to barrel or endcap M cluster. The track matcher takes the different ϕ granularities in the drift chamber (42 bins) and calorimeter (24 bins) into account.

For the drift chamber and calorimeter, the multiplicity logic defines distinct particles using variable minimum $\Delta\phi$ cuts. The multiplicity of each type of match is reported as 3-bit numbers to the global decider.

The track matcher reduces the orthogonality of the two independent Level 1 triggers and is intended for use only during very severe beam backgrounds.

Global Decider

The global decider uses the multiplicity lines from the drift chamber, calorimeter triggers, and the track matcher to issue the Level 1 strobe. It also has access to the calorimeter θ-slice number for single-prong Bhabha events that need to be prescaled as a function of θ. For a diagram of the the trigger input lines, see Figure 10-24. A completely flexible trigger is implemented through the use of a RAM-based look-up table with 24 input bits (each of which can be masked off) and one output bit. This implementation is upgradable.

10.9.5 Simulation

The Level 1 trigger was simulated using the GEANT model of the *BaBar* detector, BBSIM. For the drift chamber trigger, the scored space-point hits in each layer of the GEANT model were converted to wire hits that were fed to segment finders. The BLT algorithm [Kin89] was used to find tracks, and global trigger criteria were constructed based on counting distinct A tracks and B tracks [Kra94b]. A display of an example B^0 event is shown in Figure 10-27. A curvature finder was simulated using axial segment positions and patterns as inputs.

For the calorimeter trigger, the scored energy deposits in each CsI(Tl) block were summed into fixed 5×5 towers that were paired to form two-tower (50-block) clusters. Isolation criteria in θ and ϕ were used to count clusters [Kra94a].

The implicit transverse-momentum and solid-angle trigger acceptances for single muon tracks, identified as the A-track, B-track, and M-cluster objects of Table 10-7, are shown in Figure 10-28. Note that the inefficiency due to tracks hitting adjacent corners of four towers,

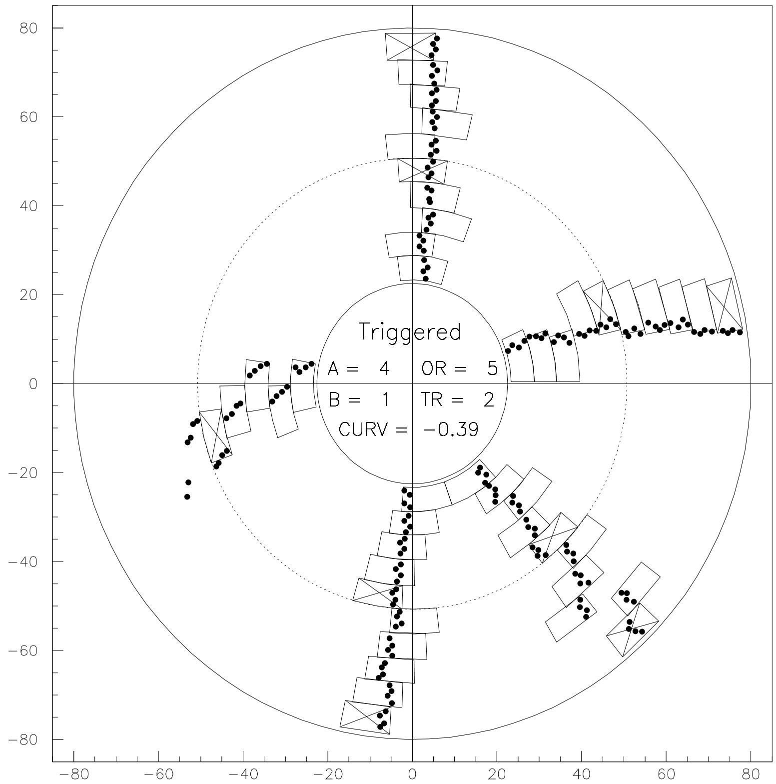

Figure 10-27. *Example of trigger response to a simulated B^0 event. The dots represent wires that were hit. Boundaries of segment-finder supercells that were hit are drawn. Crosses mark ϕ locations of A tracks and B tracks in counting layers. The event satisfied the open trigger requirements.*

thus depositing only half their energy in one domino, is neglible. Event efficiencies and rates using this simulation are given in Section 10.8.5.

10.9.6 Trigger System

Timing Considerations

The Level 1 trigger has a latency of 9.5 μs and a jitter window of ± 0.5 μs. These constraints are dictated by the characteristics of the front-end systems.

The latency time was set to a number that would not change and that would provide the maximum time to perform the Level 1 trigger while not degrading the physics of the vertex

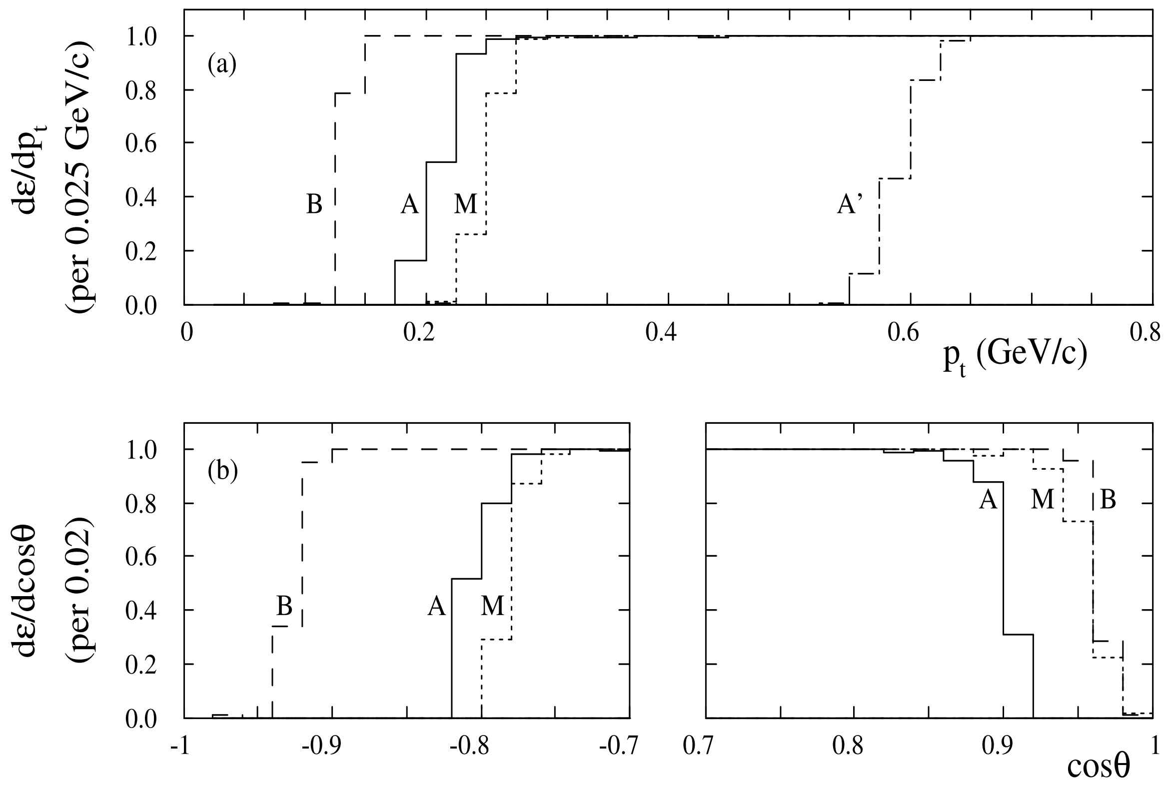

Figure 10-28. *Differential efficiencies for single muons to be identified as trigger objects A, A', B or M (see Table 10-7), as a function of (a) p_t with $|\cos\theta\|<0.7$ and (b) $\cos\theta$ with p_t >0.4 GeV/c.*

detector. At full luminosity and 10 times background, there is little impact on the physics by the latency. At full luminosity and 10 times background, the time uncertainty will result in 30% occupancy in some sections of the vertex detector due to lost beam particles.

The pipeline length is composed of several elements, listed in Table 10-12. The trigger contains circuitry to align the track linker output to the late-arriving cluster finder output before these reach either the track matcher or global decider. The estimated minimum pipeline length of 6 μs is well within the fixed latency of 9.5 μs. Part of the safety margin is expected to be consumed as detailed designs are developed for the trigger. To make up any difference or account for potentially shorter calorimeter shaping times during high-background running, the output of the global decider passes through an adjustable-delay unit before producing the Level 1 strobe.

The jitter is set by the characteristics of the drift chamber and calorimeter. Calorimeter signals will have a 1–2 μs peaking time. These signals are sampled every 0.25 μs to form the trigger primitives. Simple peak finding is used in the Level 1 trigger. This results in

Path	Latency
Detector Signal Development	
Drift chamber Drift Time (600 ns)	
Calorimeter Peaking Time (4000 ns)	4000 ns
Transmission to readout modules	100 ns
Logic to Form Local Trigger Primitives	
Drift chamber Segment Finder (250 ns)	
Calorimeter Tower Summer (250 ns)	250 ns
Transmission to Particle Finders	
Drift chamber Track Linker (300 ns)	
Calorimeter Cluster Finder (250 ns to 450 ns)	450 ns
Logic to Form Trigger Decisions	500 ns
Transmission to Fast Control Distribution	100 ns
Transmission to Readout Modules	100 ns
Transmission to Vertex Detector Front-End Cards	100 ns
Total Minimum Pipeline Length	5600 ns

Table 10-12. *Level 1 trigger pipeline length.*

finding the peak to within one time bucket or a resulting jitter of $\pm 0.25\,\mu$s. Drift chamber signals are sampled at a faster rate. They are extended by digital one-shots of 600 ns (the maximum drift time). Decisions are made every 250 ns by the segment finders. The resultant jitter window is 850 ns. A trigger jitter shorter than $\pm 0.5\,\mu$s would reduce the storage and bandwidth requirements on the front-end systems, which could be significant for the vertex detector.

Globally, these triggers from the drift chamber and calorimeter are processed to form a global trigger by the track matcher and global decision maker every $0.5\,\mu$s. This specification is set to be consistent with the Level 1 trigger jitter requirement of $\pm 0.5\,\mu$s.

System Functions

The Level 1 trigger system interacts with almost all other components of the *BABAR* electronics and online computing systems.

Global triggers will have prescaling of the selected trigger types. This prescaling is included to allow for optimization of the desired rate for physics data and for monitoring functions such as Bhabhas.

It is necessary to have timing and control interfaces to the fast control system. The trigger delivers the Level 1 accept signal to the fast control system. Fast control produces the Level 1 strobe for the front-end systems. The trigger is a consumer of fast control signals like all other front-end systems.

Monitoring of deadtime and efficiency must be part of the functionality of the electronics system. Deadtime monitoring is accomplished in the fast control system by measuring the physical time the trigger is not accepting new triggers and monitoring of scaled-down triggers such as Bhabhas. The efficiency of the trigger will be determined by the online system that is monitoring the primitives and resulting triggers. The scheme to perform this functionality must be included in the online design.

Readout by the data acquisition system will include all trigger primitives, drift chamber and calorimeter trigger output, and global decisions. This is needed to monitor the efficiency and operation of the trigger for both online and offline analysis.

The trigger will be partitionable such that it can run at the crate level, trigger system level, and *BABAR* system level. This functionality is included to allow for installation, debugging, and system testing. As part of the testing, the system will have the capability of injecting a calibration signal into the front-end electronics that allows for the determination of trigger primitive thresholds. Provisions will also be made for memories that can be downloaded with patterns which can be played through the trigger. These functions will facilitate the characterization of the trigger and help determine whether it is operating as desired.

10.10 Data Acquisition

10.10.1 Introduction

The key function of the *BABAR* data acquisition system is to collect the event data selected by the Level 1 trigger from the front-end and trigger systems, and transfer these data into a processor farm for packaging and filtering. Both hardware and software are required to concatenate the event fragments from all the data sources into complete events. Hardware and software protocols are needed to provide error-free transmission of the data, detect invalid operating conditions, prevent buffer overflows, and throttle the trigger rate. The data acquisition system also provides calibration, monitoring, and debugging facilities. The specialized equipment needed to calibrate the system electronics, including fast strobes, software processes, and hardware verification is a natural extension of data acquisition activities. Monitoring of front-end and trigger components, providing instantaneous status information, localization of errors, and error recovery tasks such as the flushing of queues are a joint responsibility of the data acquisition system and the online computing system.

The data acquisition system interfaces to all detector front-end systems for event data collection, to the trigger systems for both data collection (of trigger information) and event data flow control, and to the online system for setup and runtime control as well as monitoring of the data acquisition system.

Data are transferred downstream within the data acquisition system in response to the identification of events by the trigger system. Under normal conditions, the Level 1 trigger rate is expected to be less than 1.5 kHz with the open trigger shown in Table 10-9. The largest component in this rate is cosmic ray events. Under severe machine background conditions, the Level 1 trigger rate can be kept below 2.0 kHz using the safe trigger shown in Table 10-11. In subsequent trigger levels, the drift chamber information and the silicon vertex detector information are used to restrict the tracks to those that originate near the interaction point. The final Level 3 trigger is expected to send approximately 100 events per second to be logged to tape. Contributions to this data rate are produced in many crates and flow through the event-assembly network over many individual data paths. Thus the total bandwidth can be easily achieved using any one of a variety of modern transport mechanisms. The *BaBar* detector trigger and data acquisition system is fully pipelined in both the trigger and the data flow paths. As a result, the system is nearly deadtime free at background levels up to ten times the rate given by machine simulations. Commercial equipment is used whenever suitable modules are available. The architecture is such that the capacity of the system is easily upgradable with the addition of data paths and processors.

10.10.2 Requirements

The requirements for the *BaBar* data acquisition system fall into four broad categories. First are quantitative requirements such as those on the bandwidths of data paths and the number of data sources to be accommodated. Next are architectural requirements that allow front-end and trigger systems to be operated as a complete system, or allow portions of each to be operated as separate partitions. Then there are requirements for initializing, controlling, and monitoring the system. Last are quality control requirements for items such as reliability, error rates, and maintainability. For a more complete discussion, see the *BaBar* electronics requirements document [Req94].

Quantitative Requirements

The data acquisition system must satisfy the requirements for event rates, bandwidths, latencies, and deadtimes. The fundamental requirement of the data acquisition system

Detector System	Number of Readout modules	Number of Crates	Number of Data Sources
Vertex Detector	26	2	5
Drift Chamber	285	18	18
DIRC	3	1	1
Aerogel (ATC)	8	1	1
Calorimeter	280	18	18
Muon System (IFR)	7	2	1
Trigger	3	3	3
Total	612	45	47

Table 10-13. *Numbers of readout modules, crates, and data sources for each detector system.*

Maximum Trigger Rate	2 kHz
Max. Level 1 Latency	10 μs
Max. Average Event Size	25 kbytes
Max. Event Data Rate	50 Mbytes/s
Max. Deadtime	10%
Number Data Sources	47
Number Level 3 Processors in Farm	10

Table 10-14. *Requirements of the data acquisition system.*

is that it be capable of collecting all event data for those events selected by the trigger, while imposing less than a specified amount of deadtime. Table 10-13 shows a list of all the detector systems that generate event data. The middle column of the table gives the number of readout crates required to house the modules associated with each detector system; the number of data sources for each detector system is shown in the last column. Most crates contain only one data source; the crates for the vertex detector each contain two or three data sources.

Together, all sources are expected to generate approximately 25 kbytes of data per event, with a combined event data rate of 50 Mbytes/s. Other requirements are derived from the expected physics and trigger rates, and the deadtime requirement of less than 10% at design luminosity. A list of important requirements is given in Table 10-14.

Architectural Requirements

The system must scale modularly to accommodate upgrades, since it is inevitable that the data acquisition requirements will change during the life of the *BaBar* detector. A key element in this scalability is the design of the event-assembly network.

Furthermore, the data acquisition system must be partitionable. Partitioning means logically dividing the electronics of the detector into two or more independent subsystems, each with independent control. This allows multiple detector systems to perform calibration, diagnostic testing, and data acquisition tasks simultaneously and without interference. With this design, it will be possible to assign individual crates or entire subsystems to a partition under computer control without the need for recabling.

Control Services Requirements

The data acquisition system must also provide a set of services to control and utilize the calibration and monitoring facilities built into the various front-end and trigger subsystems. It must recognize and handle error conditions—*i.e.,* errors in the transmission of data or malfunctions of components of the system. It must record non-event data streams such as magnet currents, high voltages, and accelerator parameters. It must initialize, download, and control the various processors and modules distributed throughout the system. Finally, it must deliver event data to consumer processes reliably.

Quality Control Requirements

The data acquisition system must ensure the integrity of the event data stream. It must transport data from the front-end and trigger systems without errors in either data content or the identification of origin.

Care is being taken to ensure a high level of reliability. Proven industry standards and well-supported commercial software and hardware will be used wherever possible. Diagnostic features such as error detection and correction schemes will be implemented. To minimize downtime, the data acquisition system will be able to detect operational errors or malfunctioning modules and correct or report these conditions. Most system hardware will be implemented with modular components to permit the rapid replacement of malfunctioning units.

10.10.3 Architectural Overview

A block diagram showing the main components and interfaces of the data acquisition system is shown in Figure 10-29. In this figure, the data flow from left to right. The main components are the readout crates (VME), the event data flow control system (fast control), the event-assembly network, and the Level 3 event processing farm. The main interfaces shown are the front-end interface, the Level 1 trigger interface, and the slow control interface. Within each readout crate are the components of the various detector front-ends, trigger interfaces, and data acquisition system. These components collect analog data from the detector front-ends, process them in response to a trigger signal, and deliver digitized data to the downstream parts of the data acquisition system. The main components of the data acquisition system are described in the following sections.

Crate-Based Electronics Overview

Within the readout crates, there are four principal components or modules:

- readout modules that interface to the detector-mounted electronics and, for the drift chamber and calorimeter, provide specific information to the prompt trigger;
- fast control distribution modules that interface the crate to the fast control system;
- a readout controller that connects to the detector control system, provides monitoring and diagnostics, reads the event data fragments in the crate, and controls the flow of data through the buffers; and
- an event-data port card (data source) that interfaces the crate to the event-assembly network.

In each crate, the readout module connects the detector-mounted electronics to the detector-specific trigger, fast control system, and readout controller. Since most of these interfaces are identical for all detector systems, it is useful to look at the common features. Emphasizing the commonality of the systems creates a unified data acquisition structure. Although there are unique interfaces to the on-detector electronics and to the trigger, there is a common interface to the fast control system and another to the readout controller through the crate bus.

Event-Data Control Overview

The flow of the event data through the data acquisition system is controlled by the event-data control system. Its primary function is to keep data flowing efficiently from the

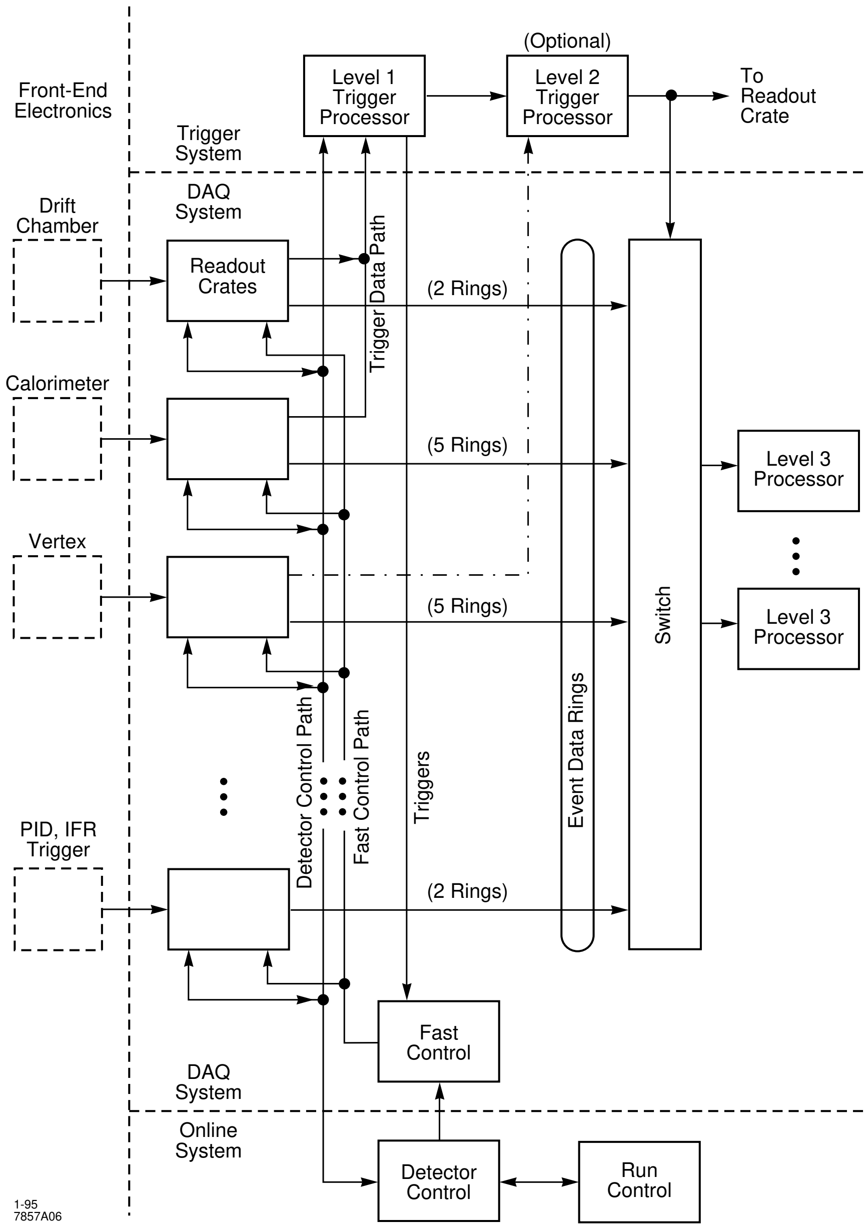

Figure 10-29. *Block diagram of the Data Acquisition system.*

front-end electronics, through the readout and trigger-system crates, through the event-assembly network, and to the Level 3 processor farm. The data acquisition system has three separate mechanisms for providing the event-data control: high speed control and trigger throttling (the fast control system); VME-based data collection and flow control algorithms; and network control and resource allocation.

The high-speed flow control uses a "backpressure" model to guarantee the orderly delivery of data. The mechanism works as follows. Ten μs after a particular beam interaction occurs, the Level 1 trigger declares a Level 1 accept. The fast control system fans out this Level 1 accept signal to all crates currently accessible to the main trigger partition. The signal is received synchronously at these readout crates by the fast control distribution modules. If this module is configured to respond to signals from the main partition, then the Level 1 accept is propagated though the crate's fast control bus to each readout module. When input queues on the readout modules become full, a busy signal is transmitted back through the fast-control system to the trigger.

When a readout controller processor has completed processing an event fragment, the data will be located in a region of memory accessible by VME. The processor will also place the location of this event fragment and the length of the fragment in a separate queue, also accessible by VME. If either the data queue or the pointer queue becomes full the processor stops and the VRAM fills, resulting in a throttle condition. The locations of the heads and tails of both queues are accessible to both VME and the processor. The read-modify-write protocol of both the VME and processor ensure that the queues are properly maintained.

Each readout controller within a partition is assigned a unique position in the readout sequence. When the readout controller in the first crate in the sequence finds that it has a block of events ready and receives an event-request token, it transmits the event-data fragment to a designated Level 3 processor. When the transmission is complete, the token is passed to the next readout controller in the sequence. This process continues until the last crate in the sequence is reached. At that point, the token is sent back to the Level 3 processor. When the complete event has been successfully read, a message is broadcast to all readout crates signaling the successful completion of the event assembly. A suitable algorithm will distribute events among the Level 3 processors.

Event Assembly Overview

The task of the event-assembly network is to send the event fragments from the 47 event-data port cards to the designated Level 3 processor.

Some of the requirements of the event-assembly network are listed in Table 10-14; these numbers will change as the design progresses. For example, the number of destination processors will depend on the computational speed of the processors, refined estimates of the computational capability required to process a single event, and refined estimates of the event rates generated by the trigger processing.

The event-assembly network will be based on a commercial network similar to those presently being used or under development for the computer and telecommunications industries. Such a network, combined with the use of standard communication protocols, has many advantages. The hardware is readily available from multiple vendors. It is reliable and has a good price/performance ratio, especially when development costs are considered. The communication protocols are already implemented and include extensive error checking, flow control, and collision management. The data are usually self-directing and do not require any central intelligence to control the data flow. The amount of programming labor that is saved by employing this solution instead of a home-grown network is enormous.

The benefits of an interconnect based on commercial networks are not free. The key disadvantages are constraints imposed on the architecture by network hardware and protocols. The advantages, however, include the possibility of implementing flexible load-balancing algorithms in software without changes to the hardware configuration and the capability of scaling the system incrementally.

The flow control and error recovery procedures of the network protocol can be used to implement a simple data-flow architecture. A distributed flow control system allows high link utilization without special hardware to implement control functions. Readout synchronization in the trigger system and front-end electronics is required to achieve this benefit. A Level 1 trigger inhibit must be asserted if any readout crate lacks sufficient buffering to read out its portion of an event. This prevents partial readout of an event.

The network protocols and interfaces have been designed for efficient transmission of large messages and modest transaction rates. Thus, a message size of about 4 kbytes or more is required for efficient use of the network. This will require fragments containing data from several events to be sent in each message originating at a data source in a readout crate. This has the advantage of decoupling the real-time event rate from the rest of the data acquisition system. Sufficient memory must be provided at appropriate points in the system to buffer the large messages.

Online Interface Overview

Software development for the readout controllers and readout module processors requires a commercial development environment and a commercial real-time operating system. The *BaBar* computing group is considering the VxWorks real-time operating system and the EPICS control system software tool kit for use in the control systems. Broad use of these tools throughout the computing, online, data acquisition, and trigger systems permits shared expertise and seamless integration of experiment control and data acquisition. The online interface provides the physical means whereby operators and applications software have the ability to control and monitor the many processors in the front-end and data acquisition systems.

10.10.4 Readout Crates

The readout crates provide housing for several types of modules required for the electronics system. These include the readout modules, the fast control distribution module, the readout controller, and the event-data port card module. The readout crates will be standard VME crates. All readout data, messages, and monitoring are to be transferred between modules on the VME backplane bus. Lines on the VME P2 bus are to be used to transmit high-speed control, trigger, and timing signals. Each crate will have spaces for 21 modules. These will be able to accommodate the readout controller, the event-data port card, the fast control distribution module, a temporary bus monitor module (for debugging), one space reserved for future use, and the 16 readout modules.

Readout Modules

All detector systems are able to present a common hardware and software interface to the central electronics system because of the design of the readout modules. As shown in Figure 10-30, each card contains a standard set of functions, which are shown as solid blocks. The "personality" section, shown in the figure as a dashed block, is designed to accommodate each detector system's special characteristics. It is expected to make provision for bi-directional data flow between a readout module and its associated front-end electronics. This will permit, for example, dummy events or configuration data to be downloaded into the front-end electronics for tests of the overall system.

The readout module receives triggers, clocks, fast resets, *etc.*, via the crate's fast control bus. Triggers are time stamped and buffered in the FIFO before being presented to the card's CPU. This feature buffers the CPU from the rate fluctuations. The FIFO *almost full* condition results in trigger throttling; receipt of a trigger in the *full* condition generates an error signal. The interface between the personality section and the readout module will convey all necessary signals to permit any special processing needs of the detector systems. These may include analog signal processing, fast digitization, and data buffering. A bit-parallel digital data bus will convey data over the interface. A serial clock signal strobes the data into or out of the VRAM depending on the state of a direction signal. The clock line stops whenever there are no incoming data. The numbers of DAQ readout cards required by the various detector systems are given in Table 10-13.

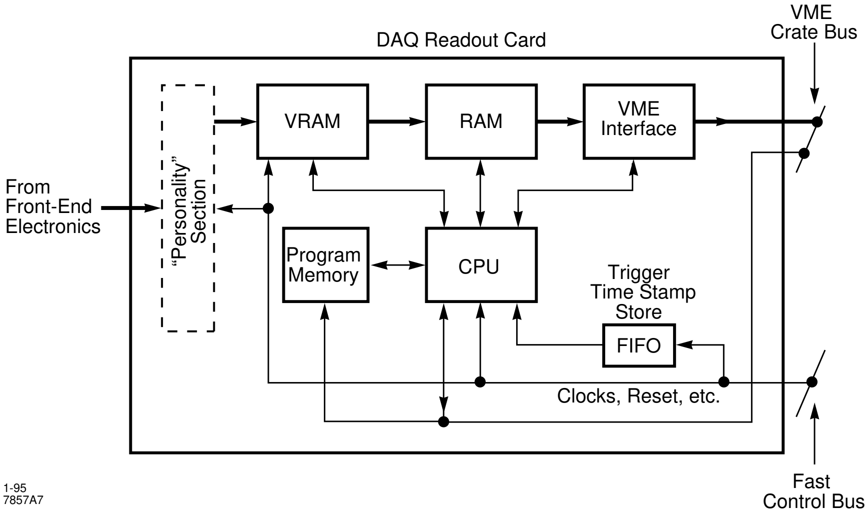

Figure 10-30. *Block diagram of a DAQ readout card.*

To achieve event synchronization across the interface, all data streams and event serial numbers are initialized by a system-wide reset signal. The read strobe then ensures that the multiple data streams remain synchronized. The personality module sends an error signal when it detects transmission errors relative to the incoming data stream.

Fast Control Distribution Module

The fast control system provides all of the high speed timing signals required for the selection, synchronization, and identification of data. These functions include fast reset, initialization, synchronization, clocking, calibration, triggering, and busy indications. The system also conveys those signals needed for rapidly enabling and disabling the triggers as required by conditions of filling and emptying of data buffers. A crate fast control bus will augment the VME backplane, and many of these signals will be transmitted over the unassigned pins of this bus.

The fast control distribution module is a VME card that receives the high-speed signals from the fast control system and repeats the signals on the crate fast control bus. This card is able either to drive directly the signals received from the clock and control interface or to generate these signals under program control through VME. This latter capability is needed to perform stand-alone operation and calibration of systems created by forming partitions. For example, the timing card can be told to issue a trigger or a calibration strobe at a specific moment, or select a particular partition with which to communicate. In addition, this card performs various control functions within the crate, including receiving a trigger throttle signal and an error signal from each readout card. These signals are then made available to the central fast control system, to be used in the trigger throttling algorithm.

Readout Controller

The readout controller is the only VME master contained within each readout crate and so is responsible for many activities, the most important of which is the event readout. In addition, the processor on this board is responsible for downloading software to all processors within the crate, responding to crate interrupts, performing detector control, monitoring, calibration, debugging, and operating in stand-alone mode. Since EPICS is likely to be used in the detector control subsystem, this processor must operate under VxWorks. While minimal memory requirements are needed to support the small event sizes of the *BaBar* data stream, excellent processor throughput will be required to assist in data formatting, queue management, and interrupt servicing.

The processor on the readout controller receives a trigger as an interrupt. After the processor services the interrupt, it executes user-defined codes to allocate buffers and read in the data. It builds a list of commands to transfer these data and then performs them as a sequence of high-speed input-output operations. When the readout controller determines that a complete event has been processed within the crate, it waits for a signal (token) before requesting transmission of the event fragment to the event-assembly network. Due to occasional long latencies in the readout, there may be several events ready for readout.

Event-Data Port Card

The event-data port card (data source) in each crate interfaces to the transport medium for all event data streams and is responsible for transmitting the crate's event fragments via this medium to the event-assembly network. Since none of the expected data streams exceeds 5 Mbytes/s, many data transmission technologies and logical arrangements are possible candidates for this task. *BaBar* uses switched FDDI.

The data sources associated with the various detector systems output data at widely different rates. This makes it quite reasonable to arrange the data sources into a topology of several rings, thereby reducing the number of virtual sources seen by the event-assembly network. This arrangement is discussed in detail in Section 10.10.6.

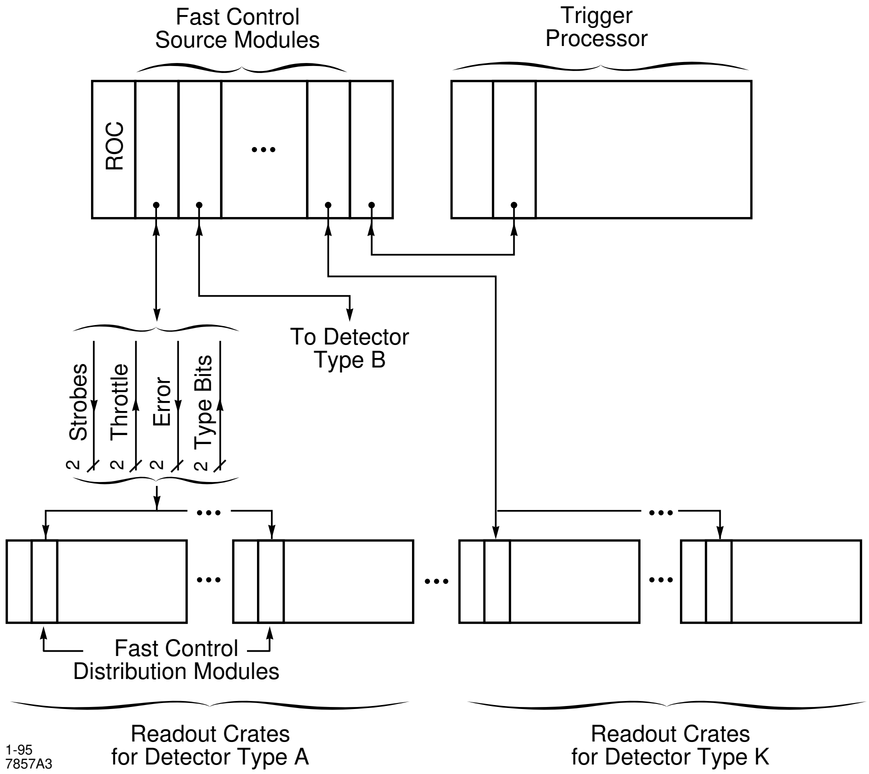

Figure 10-31. *Partitioning logistics.*

10.10.5 Event-Data Flow Control

Partitioning

It is important to be able to partition the electronics systems for the various detectors into independently operating systems for commissioning, calibration, and repairs. It is also useful to be able to run the detector in normal data-taking mode even though some parts are unavailable for various reasons, including being independently operated. Counting the trigger, *BaBar* contains seven independent systems. It will be possible to create at least seven partitions, allowing the systems to be initially commissioned independently. Figure 10-31 illustrates the partitioning scheme.

The lowest level of granularity in partitioning is the crate. A crate is either offline or online in a partition. A crate in a partition is under the direct control of the data acquisition system. An offline crate can still communicate with the online software for purposes of reading or setting voltages, running dedicated software diagnostics, *etc.*, but it is not included in the normal data acquisition operation.

The partition containing the trigger system is referred to as the main data-taking partition. From the trigger system, triggers are distributed to other systems via the fast control distribution system. We do not envision the trigger system itself as partitionable, although it may be possible to take some of its crates offline for limited repairs. A crate which is offline, or which is not in the main partition, may not be generating valid trigger primitives and will be ignored by the trigger system. This will be set up as part of the partition initialization process.

The crates in a partition other than the main partition still communicate using the fast control distribution system and the data-transfer protocols. A system within that partition may be responsible for generating strobe signals for distribution in lieu of triggers, for example, to control calibration and diagnostic activities. It is also possible to define a partition that does not contain a module to drive the fast-control system; this partition is then fully functional from a data-transfer perspective but will require that all processing be initiated by software messages instead of by hardware strobes.

Crates can be read out via the main data transfer path or at a much slower rate via the controls local-area network. Depending on the final architecture for the main path, it may be possible to divide it logically among partitions or to share it on a first-come first-served basis. Either of these options will provide acceptable functionality for commissioning and calibrating the detector.

Fast Control System

Clocks, strobes, type-bits, throttle signals, and error indications must be distributed among the data acquisition crates. These must be distributed in a flexible way to allow systems to use them for their own commissioning needs within special purpose partitions.

The free-running nature and small physical size of the *BaBar* electronics system makes timing and control distribution relatively easy. We plan to have a few central modules (fast control source modules) generating the required timing and control signals. The timing signals will be derived from the PEP-II rf clock and will maintain a stable phase relationship at the nanosecond level on the module output.

The set of strobe signals is used to carry time-specific information such as the Level 1 trigger and calibration strobes. Type-bits are used to indicate special purpose strobes, for example, an occasional diagnostic trigger which requires custom processing by the data acquisition system. Throttle signals are used to hold off further strobes when a system or crate is unable to process further requests for a short time, or they can be used to indicate that processing of a given strobe is complete. Error indications are used to inform the online system that some condition has been detected that requires higher level intervention.

The main partition will have type-bits, to classify events as, for example, normal, Bhabha, uncompacted, and snapshot. The last two are typical of several types used for occasional checks of the data acquisition logic. They will carry the Level 1 accept on a strobe and may use a second strobe to inform data acquisition modules in advance of a diagnostic trigger. Strobes in the main partition, including any that need interlocking to the throttle signal and special intertrigger time requirements. The fast control system monitors trigger deadtime.

During system-specific calibration, one might use several strobes to command the crates to generate local timing signals, and then trigger data readout on the result. It is the responsibility of the particular system to generate the required strobe signals, with interlocking to the throttle control as appropriate. If a system has only one set of electronics for generating these signals, it will not be possible to split the system into two fully functional partitions.

Fast Control Distribution Module

We currently envision eight strobes, eight busy signals, eight error signals, and eight sets of type-bits. This is enough to allow all seven major systems of the *BaBar* detector to run in independent partitions when required.

To distribute these signals, each data acquisition crate will contain a fast control distribution module. It receives the entire set of signals and provides a standard interface to the crate. It can be programmed via the VME interface to route particular strobes and/or type-bits to either front panel connections or a set of reserved backplane signals, or both. Additionally, VME interrupts can be generated from strobes. Throttle and error signals can be driven from front panel connectors or from backplane signals. All signals can also be interrogated via VME read operations for diagnostic purposes. VME-readable scalars and LEDs are provided for most functions to help in commissioning and debugging.

Fast Control Source Module

The fast control source module is used to drive the distribution system. It is configured via VME to drive strobes and type-bits from dedicated backplane signals or front panel connections, and to break out throttle and error signals onto the backplane, front panel, or into VME interrupts as required. Typically, the error signal will interrupt the local readout controller, which will poll the appropriate modules to find the cause of the problem. All signals can also be interrogated via VME read operations for diagnostic purposes. One or more of these fast control source modules is then used by the trigger or calibration control electronics in each system to distribute the required strobes, and other signals to the other crates in the system for operation in a partition. It may be possible to combine both the fast control source module and fast control distribution module onto a single VME card to reduce expense.

Using this system, the online software can flexibly allocate signals to particular partitions by controlling the settings of the control module(s). This, in combination with informing the trigger system of which crates to monitor and setting up the data-transfer network appropriately, allows each partition to function independently.

Throttling and Buffer Management

The basic hardware throttle for the *BaBar* data acquisition system is provided by the throttle lines of the fast-control system. Asserting the relevant line for a partition inhibits its strobes. In the data-taking partition, this means that distribution of the trigger is inhibited, causing deadtime. To allow an accurate determination of deadtime, there will be pairs of scalers counting system clock pulses—one gated by throttle and the other not. One pair of scalers will be assigned to each possible partition. In addition, the trigger will count Level 1 strobes internally, even when the data acquisition system is inhibited. All these scalers will be accessible as part of the normal event stream.

There are two sets of buffers that the fast control system must protect from overflow. The first of these are contained in the readout chips of the vertex detector electronics. These contain on-chip back-end buffers which are of fixed length and three events deep. They buffer against the variations in transmission time along the fiber to the readout modules. The fast-flow-control system must promptly assert throttle to prevent any overwriting of data if all three on-chip buffers on any readout chip are occupied.

Due to the limited space available, this portion of the fast control system will reside in the link controller at the data acquisition end of the line. The link controller will keep track of the chip buffer occupancy by monitoring a local model of the chip buffer states.

The other buffers controlled by the fast control system are the VRAM buffers on the readout modules. These are operated as circular buffers, in some cases (such as the drift chamber subsystem) with addresses linear in time, and in other cases, (*e.g.,* the vertex detector system) with addresses linear in event number. In the event-linear case, the throttle will be applied if there is too little space to accommodate two maximum-size events.

There is a third source of fast control signals. When a Level 1 strobe is presented to the readout module, the address in the VRAM of the corresponding data needs to be noted for later digestion by the readout module's CPU. This work queue will be maintained in a hardware FIFO. This FIFO will also be able to generate throttle if it is at risk of filling completely.

The throttle signals from each readout module will be presented individually to the fast control distribution module in its crate. This will allow disabling of these signals from broken hardware.

Event Data Flow Control

The fast flow control system is the only form of event data flow control that directly throttles the trigger. All later stages of control operate by a backpressure mechanism.

In the backpressure scheme, any stage in the event flow—starting with a processor in the Level 3 farm—which fills up simply refuses to release resources used by its upstream neighbors. This eventually causes the upstream neighbor to fill. The process cascades until the upstream neighbor is the memory on the readout module. When that memory fills, throttle is asserted and deadtime is introduced.

These algorithms become more complex in the case in which events are bundled into super events to amortize the cost of interrupts. Note that bundling events can also reduce the interrupt load within the crates; the fast control system could be programmed to provide interrupts to the readout controller only after a designated number of events is available in the readout modules.

10.10.6 Event Assembly

After the delivery of the Level 1 accept, the readout modules and readout controllers proceed autonomously to the point at which all useful event data from a crate are collected into that crate's readout controller. At this point, the data represent the crate's event fragment. The event assembly process collects all the crate-level fragments for a particular event into a single Level 3 node for further filtering, reconstruction, and storage. The baseline technology choice is switched FDDI, a commercially available high-speed network.

Such an assembly network must provide enough throughput to handle the worst-case average event size at the Level 1 accept rate with a reasonable latency. Instantaneous rates above the network throughput will have to be buffered before the network or deadtime will result. Since chance fluctuations may fill any buffer, the goal is to provide enough buffering to keep this deadtime below the few percent level at 10 times nominal background.

The pre-network buffers imply a latency for assembling events that is fluctuating and potentially large compared to the event rate. Since it is easy to provide this buffering on the readout controllers, and there are no hard real-time restrictions after the data have been

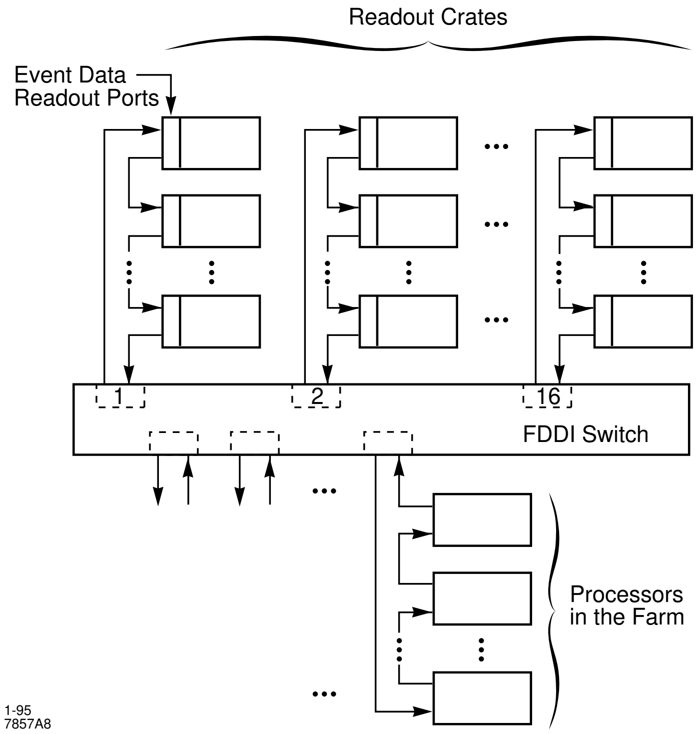

Figure 10-32. *Crate readout ring topology.*

rescued from the trigger-latency buffers, this latency is not important. The fluctuations introduced by the Level 3 nodes are likely to be much greater.

In a switched network with multiple destinations, only the switch needs to see the full data bandwidth. Typically, switches are designed with very high bandwidth backplanes. For example, one currently available FDDI switch has a capacity of over 3 Gbits/s (about 400 Mbytes/s). Real utilization approaching 80% of this figure has been observed at LBL, whereas the *BaBar* maximum demand is only 50 Mbytes/s. Hence, the switch itself should not be a limit.

After the switch, the next potential limit is the capacity of an optical fiber. At an FDDI data transfer rate of 100 Mbits/s (12.5 Mbytes/s), it is clear that a single fiber cannot cope with the load. On the other hand, assigning one fiber per data source would result in a large number of input ports on the event-assembly network, with many of the fibers significantly under-utilized in terms of data rate. A reasonable alternative is then to assign several data sources to each fiber data channel, with the number of sources per channel to be determined by the data rates of the sources.

System	# of Sources	# of Sources/Ring	# of Rings	Data Rate/Ring
Calorimeter	18	4	5	3.2 Mbytes/s
Drift Chamber	18	9	2	3.6 Mbytes/s
DIRC & ATC	2	2	1	3.8 Mbytes/s
Vertex	5	1	5	3.3 Mbytes/s
IFR & Trigger	4	4	1	3.2 Mbytes/s
Totals	47		14	

Table 10-15. *Readout ring topology and bandwidth at 2.0 kHz.*

Figure 10-32 shows a ring topology for combining several data sources or sinks onto a single fiber channel. Here, it is assumed that the event-assembly network is a 32-port device. Sixteen of the ports are assigned as input ports, with each port connected to a ring connection of data sources. The remaining ports are used as output ports to the processors in the Level 3 farm, which are also connected in rings. None of the expected maximum data rates per input ring exceed 25% of the 12.5 Mbytes/s bandwidth of switched FDDI.

These rings could be upgraded with a future higher performance technology at modest cost in order to take advantage of higher throughput. Technologies such as FCS (fiber channel standard), or SCI, can be used in a similar ring topology and provide significantly greater throughput. Technologies such as HIPPI and ATM do not have ring topologies, but suitable substitutes exist with similar performance improvements.

The sequencing of data onto the input rings is controlled by token passing. To gather event fragments, an input port of the event-assembly network will launch a token. The first data source on the ring will receive the token and will transmit its event fragment(s) via the fiber. When it is finished, or if it has no fragments to send, it immediately forward the token to the next source.

Tokens can be used to convey instructions to the data sources, *e.g.,* the Level 3 farm processor to which the current event data fragments are to be addressed.

The numbers used to assign data sources to particular rings are given in Table 10-3, where the last column shows the data rates for data sources assigned to each detector type. Table 10-15 shows how the 47 sources can be connected into 14 rings, with the rings having roughly equivalent data rates easily within the capacity of FDDI fibers.

On the output side of the switch, there needs to be at least six rings to keep the load within bounds. The number of farm nodes per output ring is essentially unlimited; *BaBar* will use whatever is needed to supply sufficient computing power.

The next potential bottleneck is the data rate capacity of the event-data port cards. These are the cards that interface a data source or a farm processor node to an FDDI optical fiber. Of course, every event-data port card will put bits on the fiber at the full 100 Mbits/s for the duration of its packet. The first question is how quickly it can or needs to supply packets so that the occupancy of the fiber is as high as possible. The second question is how much local processing power in the event-data port card is required to support a TCP/IP protocol at high rates. There are reports of event-data port card/processor combinations which can achieve full fiber loads even while running TCP/IP over the fiber, but these figures are likely to be reached only under ideal circumstances. Fortunately, doubling the number of nodes per ring directly cuts the event-data port card load in half, so there is a simple cure within this architecture. Also, this is a rapidly developing technology. *BaBar* will monitor developments and take advantage of performance improvements and cost reductions.

Note that, for the above minimal output configuration, the traffic of events out of Level 3, which are recorded on tape, represents less than 7 Mbits/s (800 kbytes/s). This is a sufficiently small load that it will not be discussed further.

The data rate per data source listed in Table 10-3 is the rate each crate on the input rings must handle. For some cases, such as the drift chamber, the load per event-data port card is quite modest. On these rings, the efficiency with which the FDDI protocol will multiplex the packets is of primary importance. On other rings, such as the trigger and vertex detector, the load per event-data port card is a significant fraction of the fiber bandwidth. Again, published reports indicate that such a throughput is attainable, but much depends on the event-data port card/processor combination; the *BaBar* readout controllers need to do more than simply feed the FDDI event-data port cards. Measurements on prototype configurations and modeling of more complete configurations will be done to decide if this is a serious problem.

Because the maximum throughput of the above configuration is determined by the input fibers, there is a simple chain of consequences. First, the input fibers must be used as efficiently as possible. Second, to use the fiber efficiently, each event-data port card should be ready at any time to send useful data. Third, for data to have a useful data density, the ratio of data to overhead should be high.

The second consequence suggests maximizing concurrence in the data gathering process both within events and across events. In particular, one would like each crate to send its portion of a given event to the destination farm node without waiting for other crates. In the FDDI protocol, if two nodes try to send to the same destination, one of them will be held off until the other finishes. Within a given input ring, this is not necessarily bad. If all crates tried to send at once, they would indeed serialize, but the fiber would be fully occupied. Unfortunately, however, the entire drift chamber fiber, for example, would have to be idle while the vertex detector crates on another ring are transmitting data. The solution is to have one ring's crates transmit a different event to a different destination while another ring's crates are active. Hence, multiple events will be gathered simultaneously. Other data transmission technologies may not have this problem.

Because of contention for the output fiber bandwidth, some form of traffic shaping or control of fragment assembly order will improve event gathering throughput. An example of such an algorithm has been presented in Reference [Dou94]. In this scheme, the data acquisition crates are logically arranged in a ring. A farm node volunteers to handle an upcoming event by sending a token to the first crate in the ring. When the specified event arrives, the first crate sends its data to the farm node and forwards the token to the next crate in the ring. This crate sends its data and passes on the token. When the token leaves the last crate, the event is completely assembled. At no time is more than one crate talking to the destination node. Since each available farm node has registered a node with the first crate, that crate can begin delivering data for the next event before the first token has fully traversed the ring. Since the destinations of data for different tokens are different nodes, there is no contention for output bandwidth. Many variations on this idea are possible, and some may have better performance. For example, in the version outlined here, tokens do not overtake each other. This means that several events can be stalled waiting for the delivery of data from a particularly full crate. We intend to do simulation studies to evaluate the impact of such effects and ways around them.

Step three in the consequence chain is relevant since some crates, such as the drift chamber crates, have a low data rate. There is a risk that the useful data on the fiber will be swamped with headers and protocol messages. Fortunately, it is a simple matter to cluster multiple events together into bundles which are delivered as a unit into a single destination farm node. This amortizes message overhead over many events. The cost is more complex software in the farm nodes, which must deal with noncontiguous event fragments. Since the events are in random-access memory, a properly designed algorithm should minimize the performance loss. Any high rejection filters that run in Level 3 will be designed to run before the events are unscrambled. The cost of memory is such that bundles of 64 or even more events are practical.

10.10.7 Event Distribution

Once events have been fully filtered and reconstructed in the Level 3 farm nodes, they must be collected and stored on permanent media. The *BaBar* baseline design allows for both multiple partitions running simultaneously and multiple data streams within a single partition. This leads to several issues regarding event distribution.

Because crates are not partitionable, event distribution in a running partition involves little more than being sure that the participant at each stage—readout controller, Level 2, Level 3, and event consumer—knows where to forward its events. A resource-manager database will be used to control configuration of partitions to prevent stages from being assigned to more than one partition. The resource-manager database will also be a source of current-configuration information for the control-source and distribution modules to ensure sensible configurations of strobes. This database can be queried by the event-recorder consumer at the beginning of a run to record the detailed configuration of the partition for use in later analysis.

Level 3 is made up of identical autonomous processors, which could easily be distributed across partitions. While *BaBar* will retain this ability, it is not likely to be used often; the full Level 3 farm will generally be assigned to the main partition. Partitions running in local mode will typically contain a few crates directly feeding a conventional workstation running an event consumer for monitoring and analysis. Such a workstation will run a single-node version of the Level 3 software to gather data from the partition's crates and supply events to the event consumer as described below. These event consumers could be located at the institutions responsible for the systems.

Partitions will be defined and configured through an interactive interface to the resource-manager database, with some common groupings of partitionable components such as the whole of Level 3 being predefined. Running partitions may not be reconfigured. Resolution of contention for resources, such as Level 3 nodes, will be achieved by discussion among the people on shift.

Farm Node Allocation

Closely related to the problem of assigning fibers on the output side of the switch is the problem of assigning Level 3 processors to events. In CDF, for example, this decision is made by a central buffer manager which keeps track, via message exchange, of the availability of each farm node. The nodes are then queued for reassignment using a FIFO queue discipline. The disadvantage of this scheme is the high message load placed on the single buffer manager. An intrinsically distributed mechanism is needed to control the message load on any particular node.

The algorithm of Reference [Dou94], described above, may apply. This node allocation algorithm is presented as a near-round-robin queuing discipline for the farm nodes under the assumption that most of the time most of the nodes claim an event as the token passes it. Unfortunately, this will not be the case in general. First, note that the token must continue to circulate at all times. Under a lightly loaded system, most nodes will have already claimed an event which has not yet occurred and will therefore have no resources available; hence they will all pass the token. Under a heavy load, most nodes will be busy with their events and will again pass the token. In both cases, there is no speed control on the rotation of the token, and the token-related messages will increase rapidly in number. By recasting this allocation problem in slightly different terms, one can choose from a rich literature of distributed algorithms [Lam74, Ric81, Dol82, Pet82].

10.10.8 Data Integrity

The data acquisition scheme is quite complex, with data fragments being bundled and sent through a network, and the bundles of fragments assembled into events. To be sure that events are being assembled correctly, each fragment will need an event label and content descriptor. As the events are assembled, these will need to be checked to ensure that all pieces of the event are completely and correctly included. A global process will need to track all triggers and verify that they are either rejected for cause or archived.

10.10.9 Research and Development

System Modeling

Simulations are being used to refine our knowledge of the data acquisition requirements. The trigger data-flows have been studied using our most detailed simulations of the expected backgrounds. The buffering required in the vertex detector [Lev94b] and drift chamber front-ends has also been examined using these background simulations.

The task for the immediate future is to simulate the data acquisition system as a whole to investigate signaling and protocol overheads, and to ensure that the bandwidth and latency requirements are understood. We intend to do this using commercially available simulation packages and to build up a family of detailed simulations of the various parts of the *BaBar* detector and its response to the expected backgrounds. The detailed simulations of the trigger and individual systems will also continue to be improved as more detail about their functions becomes available.

Alternative Network Technologies

Asynchronous Transfer Mode (ATM) networks will allow for data signaling rates per fiber of 155 Mbits/s (120 Mbits/s usable) in the near future and have the potential for 620 Mbits/s (480 Mbits/s usable). There is very widespread industrial support for ATM, which should ensure the availability of components at reasonable cost. There are some possible problems with use of ATM for data acquisition. The Fiber Channel runs at data signalling rates from 132 Mbits/s up to 1 Gbits/s. It has fairly widespread support in industry, especially in connection with high speed peripherals. At present, there are only a few manufacturers of

Fiber Channel switches, and cost per node is relatively high. Scalable Coherent Interface (SCI) networks operate at a data signaling rate of up to 1 Gbits/s. Industry support for SCI is modest.

10.11 Level 2 Trigger

The baseline design omits a Level 2 trigger stage and builds all events accepted by Level 1. Since having a Level 2 may prove to be desirable and even necessary, we describe this option here. The Level 2 process for an event begins with the Level 1 accept. It receives input data from the Level 1 trigger and some of the readout modules. A positive decision results in an event being built. Level 2 provides an opportunity for intermediate event filters which reduce the number of events built, and thus diminishes the burden on the readout module processors, the event assembly network, and the Level 3 processors. If all envisioned cuts are made, the rate of event building could drop from the 2.0 kHz of the Level 1 accept to 200 Hz.

10.11.1 Filter

The primary candidate for a Level 2 filter is a cut on the longitudinal component of the distance of closest approach of tracks to the interaction point (a z cut), for which the primary source of information for determining this quantity is the vertex detector. The volume of data needed for a vertex detector based z cut makes gathering the required information impractical before the Level 1 strobe.

A Level 2 trigger might also implement the cut on track transverse momenta (a $p_{\rm t}$ cut), thus simplifying Level 1. At a minimum, the $p_{\rm t}$ cut requires the Level 1 trigger information from the drift chamber. It could be accomplished in hardware or software, although for some of the proposed drift chamber geometries, a hardware implementation of the $p_{\rm t}$ cut may be complex and require more latency than the 10 μs allowed for the Level 1 accept. A software implementation could benefit from additional information from the drift chamber.

Simulations of background processes indicate that appropriate cuts in z and $p_{\rm t}$ can reduce the trigger rate to near the logging rate with no substantial reduction in efficiencies, even for low-multiplicity tau decays [Lev94d].

10.11.2 Implementation

Logical

The Level 2 design should have minimum impact on the design of the rest of the system and should require no substantial hardware developments. As an example of a possible implementation of Level 2, a network with a low-overhead fast protocol can be used to send one message per Level 1 accept to the crates serving the vertex detector, the drift chamber, and the Level 1 trigger. These crates return information to a Level 2 processor, which makes a decision before feature extraction takes place.

Using multiple processors for Level 2 decouples throughput from latency. The Level 1 accept rate (between 1 and 10 kHz), combined with a reasonable limit on the number of processors devoted to the Level 2 decision making (such as 10), implies an average latency of no more than 1–10 ms. This is the same order of magnitude as the length of buffers implemented with VRAM. If the Level 2 decision can be made in time to copy the relevant data from the VRAM, movement of the data upon a Level 1 accept can be avoided, thus reducing the load on the readout module processors.

Once a Level 2 accept has been issued, event building proceeds as described in Section 10.10, but at a much reduced rate. Events could be bundled if desired. The leading engineering constraints on implementations of Level 2 include:

- the rate of Level 1 triggers;
- the network latencies for message passing;
- the network bandwidths for data transport;
- the processing overhead incurred by the network protocol;
- the number of processors sharing the loads of filtering and event building;
- the execution times of their tasks; and
- the worst-case latencies for response to asynchronous messages (typically much different for workstations and embedded processors).

Hardware

An example, based on two recent products, illustrates how a Level 2 trigger could be realized. Commercial reflective memory modules with interrupt capability have appeared in the PCI

Mezzanine Card (PMC) format. Reflective memories modules are connected together by a dedicated network. A message-passing mechanism, transparent to the user, keeps the contents of appropriately mapped locations in all modules the same. From the processors' point of view (PMC), they are accessing the same memory. VME PowerPC processors with PCI mezzanine slots are now on the market. This example uses several of these processors, each with a PMC reflective memory node. These appear as the readout controllers in the 20 drift chamber crates, the four vertex detector crates, and the Level 1 trigger crate, as ten Level 2 processors, and as the global controller. All reflective memory nodes form one logical network.

In response to the Level 1 accept signal, the global controller moves the Level 1 trigger primitives and time stamp into the next available buffer, then sends a message over the reflective memory network to the next available Level 2 processor, pointing to this buffer. The Level 2 processor then sends a request message to all the readout controllers on the reflective memory network for the Level 2 trigger symbols for the time interval specified by the stamp. Each readout controller then manages this request in its crate and responds, also over reflective memory. When all readout controllers have responded, the Level 2 processor completes the $p_{\rm t}$ and z cuts on Level 1 events.

Level 2 notifies the global controller of its decision and makes itself available for another event. For accepted events, the global controller notifies all readout modules, via their readout controllers, to save all data from the event, and then manages the remainder of the event processing in the same manner as in the baseline design.

10.11.3 Conclusion

A data acquisition system with a Level 2 filter would have greater architectural and data flow complexity than the baseline option. This may be offset by having a simpler Level 1 trigger and a reduced load on the event-assembly network and the processors. Furthermore, it may be required by the running conditions. We will continue to investigate its design, including prototyping the communication protocols and simulating the overall system, and we will keep its requirements in mind as we construct the event-assembly network. Simultaneously, we will track developments in networking technology, identify promising options, and select the most cost-effective low-risk implementation.

10.12 Global Support Electronics

10.12.1 Detector Monitoring and Control

Purpose

A primary tool for assuring the quality of data collected by *BaBar* is the careful control and monitoring of the detector hardware. This includes such traditional items as setting and monitoring power supplies and monitoring environmental factors such as temperature and humidity. It also includes monitoring to ensure the safety of both personnel and equipment, and communicating with the run control system to prevent it from taking actions which are not allowed by the current hardware status.

Design Strategy

The detector control includes areas of responsibility belonging both to the individual detector systems and to the central data acquisition group. The central group will provide a standard framework consisting of a set of displays, a set of procedures for setting and reading values from the hardware devices, and a messaging system for communication. The detector systems have the responsibility to identify their needs early in the design cycle and to provide the necessary interfaces to the standard hardware and software. The design process will use an object-oriented approach to abstract those parts of the implementation that are common to many systems, and from this, the central data acquisition group will provide a common framework. An important point is that the system designers think about these functions from the start so that the elements of the common framework can be identified as early as possible.

Among the quantities that will be routinely monitored for all detector systems are:

- all power supply voltages at the supply and as close to the point of use as practical;
- power supply currents where appropriate;
- power supply and crate temperatures;
- temperatures at sensitive points in the detector and in the electronics house;
- all gas pressures; and
- many additional items particular to individual detector systems.

The distribution of which functions are provided by software and which are provided by hardware will depend largely on the time-critical nature of that function. This will emerge as the design progresses.

Implementation

The detector control system will be implemented using EPICS/VxWorks. This modern, modular approach frees the designer to choose the hardware technology which is most appropriate to the problem and most cost effective. For example, devices which communicate via GPIB, VME, RS232, or other means can all be accommodated easily by using the appropriate software driver. The effects of changing a hardware module are thus highly localized.

To eliminate interference with data gathering, the communication path for detector control will be kept separate from that used to collect events. An Ethernet-based system using the TCP/IP protocol will be used for the backbone. Processors on this backbone in turn communicate with hardware modules via the appropriate bus, *e.g.,* GPIB, VME, or RS232. These backbone processors will run VxWorks and are seen by EPICS as I/O Controllers (IOCs).

A number of software tools will be needed. These are discussed more fully in Chapter 11. Some of the tools needed for this task are:

- A set of standard displays such as strip chart plots, histograms, and status displays. These allow the operator to keep track of the hardware and are provided by the Display Manager in EPICS.
- A messaging system for communicating with the operator and the run control system. Messages will be classified according to hazard level as alarms which need immediate action, warnings, or information. Communication with the run control system will allow it to take appropriate automatic action for alarms and warnings and to allow it to check the status of appropriate hardware before allowing a change of the state of the experiment; *e.g.,* the high voltages must be on in order for data to be taken. The run control system is discussed more fully in Chapter 11.
- An interface to the online database, which is needed in order to store target values, tolerances, and measured values.

10.12.2 Data Monitoring Support

Monitoring is the continuous watching of the incoming data and the hardware to spot problems as soon as possible. It is a front-line tool to ensure data quality. Examples include synchronization checks, comparing histograms of raw data to expected distributions, and monitoring event sizes. Monitoring of the detector control hardware is discussed in Section 11.

Monitoring of the physics event stream is primarily a software task and is discussed in Chapter 11. An electronics implication is that data acquisition hardware may need to supply a buffer in which sampled events are stored in such a way that they do not interfere with the primary data gathering path. Individual monitoring processes could then access this data for analysis.

10.12.3 Diagnostic Support

Purpose

Diagnostic actions are taken to isolate a problem once it has been detected by monitoring or to perform routine checks to verify that the hardware is functioning properly. They involve operations that are not normally part of data taking. Examples include executing scope loops and measuring response to known stimuli such as test pulses.

Design

The choice of diagnostic operations for a particular detector system is best left to the designers of the system, but it is important that they address the issue early in the design process. Global support from the electronics and online groups will be in the form of a framework upon which the system designers can build. This framework will come from a global, object-oriented design of the needs of the systems and will consider hardware and software together. The design methodology will be iterative to allow information learned in late stages to be fed back into earlier stages. In many cases, the choice of using hardware or software for a given function is primarily a matter of the response time needed. By deferring this decision, we can optimize before making costly decisions.

Implementation

Diagnostic runs can take place in either global or local mode. In global mode, there is central management for run control, trigger, data gathering, and timing. There can be only one such run in progress at a time. Global mode is used when the nature of the problem is such that two or more systems must run together. An example of a global diagnostic run is a pattern load of data into the system buffers to verify that the data are being read out correctly. In local mode, the system is in a partition that has its own local run control, local readout strobe, local data gathering, and local timing signals. This will be implemented by providing a local timing controller and a local readout controller for each partition. In global running, these units act under the control of a central partition. In local running, the timing controller can either pass on signals from the central partition, *e.g.,* clocks and synchronization pulses, or use locally generated ones. Similarly, the local readout controller can either pass the data on to the central partition or handle it locally. Many partitions can be running independently and concurrently, and within a given partition, it may be possible to have multiple concurrent activities. The local partitions must have access to the central database.

10.12.4 Calibration Support

Purpose

Calibration involves determining the parameters needed to turn digitized data like ADC and TDC counts into physical data such as energies and coordinates. This involves applying known stimuli to the analog portion of the data acquisition chain and observing the response through the digitization process.

Design and Implementation

In terms of data acquisition properties, calibration runs are very similar to diagnostic runs. As such, the general framework for detector diagnostics described above applies. The choices of how to calibrate and how often to do it are primarily up to each detector system and are discussed in their respective sections of this chapter.

Some general principles of calibration runs are: they should use the same analog and digitization electronics as the physics data; they should strive to measure one number instead of piecing together multiple measurements; and the absolute calibration may in many cases have to come from the physics data, *e.g.,* getting the absolute energy scale for the calorimeter from the Bhabha events.

Calibration runs can be in either global or local mode. An example of a global calibration run is using cosmic rays to do tracking alignments and to measure the dE/dx values from muons. An example of a local calibration run would be using calibrated electronics to inject known signals into the detector data acquisition system.

A special case of calibration involves automatic steps that need to be taken at specific times; *e.g.,* pedestal measurements which are taken at the start of each run. These processes will be fully automated.

References

[Amb94] M. Ambrosio *et al.,* "Performances of RPC Detectors with Tracking & Timing Electronics," *Nucl. Instr. Methods* **A344**, 350 (1994).

[Alo94a] A. Aloisio *et al.,* "High Density Zero Suppressor Card for the Muon Trigger with RPC for the L3 Experiment," IEEE–NS **41**, 225 (1994).

[Alo94b] A. Aloisio *et al.,* "The Suppressor Card for the Muon Trigger with RPC for the L3 Experiment," INFN–TC–94–11 (1994).

[Alo94c] A. Aloisio *et al.,* "The RPC Trigger System for the L3 Forward-Backward Muon Detector," in Proceedings of the CHEP94 Conference, San Francisco, California, LBL–35822 (1994).

[Alo94d] A. Aloisio *et al.,* "The L3 Forward-Backward Muon RPC Trigger System," submitted to *Nucl. Instr. Methods* (1994).

[Alo94e] A. Aloisio *et al.,* "Performances of the RPC Trigger System of the L3 Forward-Backward Muon Spectrometer," submitted to *Nucl. Phys.* B (1994).

[Bab94] *Letter of Intent*, SLAC–443 (1994).

[Bag94] R. Baggs *et al., Nucl. Instr. Methods* **A344**, 547–557 (1994).

[Bau92] D. Bauer, "The Impact of Two-photon Physics on a B Factory Detector," *BaBar Note* # 77 (1994).

[Bol91] T. Bolton *et al.,* "Report of the Trigger and Data Acquisition Group," in Proceedings of the Workshop on Physics and Detector Issues for a High-luminosity Asymmetric B Factory at SLAC, SLAC–373, 597 (1991).

[Boz93] W. Bozzoli *et al.,* "Data Transfer and Distribution at 70 Mbytes/s," CERN/ECP 93–7, (1993).

[Buo] S. Buono *et al.,* "Studies on RIO/HiPPI Based Event Building," RD13 Note N.70 (1993).

[CLE91] C. Bebek *et al.* (CLEO Collaboration), "CLEO-II Trigger System," *Nucl. Instr. Methods* **A302**, 261–276 (1991).

[Cul91] D. Cullen-Vidal *et al.,* "D0 Level-2/Data Acquisition; the New Generation," in Proceedings of the 1991 International Conference on Computing in High-Energy Physics, 659 (1992).

[Dol82] D. Dolev, M. Klawe, and M. Rodeh, "An O(n log n) Unidirectional Distributed Algorithm for Extrema Finding in a Circle," J. Algorithms **3**, 245–260 (1982).

[Dou94] D. Doughty Jr., D. Game, L. Mitchell, G. Heyes, and W.A. Watson III, "Event Building Using an ATM Switching Network in the CLAS Detector at CEBAF," Presented at the International Data Acquisition Conference on Event Building and Event Data Readout, Batavia, Illinois (1994).

[Eis90] A. Eisner, "Bhabhas at an Asymmetric B Factory," *BaBar Note* # 32 (1990).

[Ell92] N. Ellis *et al.,* "A Calorimeter-Based Level-One Electromagnetic Cluster Trigger for LHC," in Proceedings of the 1992 International Conference on Computing in High-Energy Physics, edited by C. Verkerk and W. Wojcik (Annecy, France, 1992), 210, (1992).

[Gen93] P. Bailly and J.-F. Genat, "A 100 Picosecond Resolution, 6 Microsecond Full Scale Multihit Time Encoder in CMOS Technology," in Proceedings of the Third Int. Conf. on Electronics for Future Colliders, Chestnut Ridge, p. 57 (1993). See also: J.-F. Genat, "High Resolution Digital Time to Digital Converters," in Proceedings of the First Int. Conf. on Electronics for Future Colliders, Chestnut Ridge (1991).

[Hal94] G. Haller, D. Freytag, and J. Hoeflich, "Proposal for an Electronics System for the *BaBar* CsI Calorimeter," *BaBar Note* # 184 (1994).

[Kin89] K. Kinoshita, "A Fast Hardware Trigger for Charged Particles in Multilayer Tracking Devices," *Nucl. Instr. Methods* **A276**, 242 (1989).

[Kra94a] F. Kral, "GEANT Simulation Results for a Fast Calorimeter Trigger," *BaBar Note* # 133 (1994).

[Kra94b] F. Kral and D. Worledge, "Initial Design and GEANT Simulation of a Fast Drift Chamber Trigger," *BaBar Note* # 162 (1994).

[Kra94c] F. Kral, "Simulation of Trigger Rates due to Beam-Gas Backgrounds," *BaBar Note* # 165 (1994).

[Kra94d] F. Kral, "Trigger Rates due to Cosmic Ray Muons," *BaBar Note* # 164 (1994).

[Lam74] L. Lamport, "A New Solution of Dijkstra's Concurrent Programming Problem," Comm. ACM **17(8)**, 453 (1974).

[Lam78] L. Lamport, "Time, Clocks, and the Ordering of Events in a Distributed System," Comm. ACM **21(7)**, 558 (1978).

[Lev94a] M. Levi and F. Kral, "Simulated Background Rates in Calorimeter Trigger Towers," *BaBar Note* # 134 (1994).

[Lev94b] M. Levi, "Impact of Backgrounds on Silicon Vertex Detector Architecture and Detector Trigger," *BaBar Note* # 136 (1994).

[Lev94c] M. Levi, "Obituary for the Prompt Silicon Trigger," Netnews Item *slac.b-factory.vertex # 54*, (1994).

[Lev94d] M. Levi, "Efficacy of a Level 2 Trigger Algorithm," Netnews Item *slac.b-factory.vertex # 55*, (1994).

[Lev94e] M. Levi, "Impact of the Electroproduction of Hadrons on the Trigger," *BaBar Note* # 191 (1994).

[Pat94] J. Patrick *et al.,* "The CDF Ultranet-Based Data Acquisition System," in Proceedings of the 1994 Int. Conference on Computing in High-Energy Physics, in preparation.

[Pet82] G.L. Peterson, "An $\mathcal{O}(n \log n)$ Unidirectional Algorithm for the Circular Extrema Problem," ACM Trans. on Programming, Languages and Systems **4(4)**, 758 (1982).

[Req94] "*BaBar* Electronics Requirements," in preparation.

[Ric81] G. Richart, and A.K. Agrawala, "An Optimal Algorithm for Mutual Exclusion in Computer Networks," Comm. ACM **24(1)**, 9, (1981).

[SLA93] "Status Report on the Design of a Detector for the Study of *CP* Violation at PEP-II at SLAC," SLAC–419 (1993).

[Sny92] A. Snyder, "Backgrounds in the *BaBar* Drift Chamber," *BaBar Note* # 88 (1992).

[Sny94] A. Snyder, "Electro-Production Triggers in *BaBar*," *BaBar Note* # 180 (1994).

[Wun94] S. Wunduke, M. Levi, and F. Kral, "Architecture for Readout Electronics," *BaBar Note* # 138 (1994).