DØnote # ????

A Silicon Track Trigger for the DØ Experiment in Run II

— Technical Design Report —

H. Evans[1], U. Heintz[2,3], T. Heuring[5], J. Hobbs[4], M. Johnson[2], S. Mani[6], M. Narain[2], F. Stichelbaut[4], H. Wahl[5]

[1] *Columbia University*
[2] *Fermi National Accelerator Laboratory*
[3] *Boston University*
[4] *State University of New York at Stony Brook*
[5] *Florida State University*
[6] *University of California, Davis*

March 11, 1998

Contents

1 Introduction

This document describes the design of a Silicon Track Trigger preprocessor (STTpp) for the Run II DØ detector. The STTpp is a part of the trigger system of the DØ detector and performs a precise reconstruction of charged particle tracks in the Central Fiber Tracker (CFT) and the Silicon Microstrip Tracker (SMT). It allows the selection of events that contain decays of long-lived particles, *e.g.* b quarks, at the trigger level before the data are written to a permanent storage medium, thus increasing the rate at which such events can be acquired.

The following section motivate the need to study events that contain b quarks in order to elucidate some of the most prominent questions in particle physics. Sections 3 and 4 describe the detector systems that the STTpp interfaces with. Section 5 gives a functional description of the STTpp and quantifies the expected benefits. Sections 6–8 describe the hardware implementation and sections 9 and 10 the communication with the DØ trigger system. Section 14 describes a simulation of the performance of the device. Section 15 finally gives a preliminary cost estimate and schedule.

2 Physics Motivation

2.1 The Standard Model

High-energy physics seeks to elucidate the fundamental interactions of matter. Its ultimate goal is to describe all the fundamental interactions by a single, unified theory. The standard model [1] provides such a unified description for the electromagnetic and the weak interactions. According to this model they are both manifestations of the same interaction. A phenomenon called electroweak symmetry breaking makes them appear as two very different interactions at low energies. However, at high energies the distinction blurs. The hope is that at higher energies, the distinction between the electroweak and the strong interaction, described by Quantum Chromodynamics (QCD), will disappear, and at yet higher energies all four fundamental interactions can be unified into one.

The heart of the Standard Model is the mechanism that breaks the symmetry between the electromagnetic and the weak interactions. This mechanism explains how the particles mediating the weak interactions, the $W^{\pm}$ and Z^0 bosons, can have large masses, while the photon, which mediates the electromagnetic interaction, is massless. The large masses of the W and Z bosons are responsible for the short range and weakness of the weak interactions. This mechanism also predicts the existence of a neutral scalar particle, the Higgs boson, which has not been observed yet. The masses of the fermions (leptons and quarks), which make up the matter surrounding us, are described by their coupling to the Higgs boson. However, while the masses of the W and Z bosons can be precisely predicted by the model, the fermion masses are free parameters. This lack of predictiveness is one of the major shortcomings of the Standard Model.

At the energies probed by past experiments, the Standard Model gives a surprisingly accurate description of the phenomenology of the electroweak interactions. However, if there is unification at higher energies the Standard Model cannot be valid up to arbitrarily high energies. Many possible extensions to the Standard Model, which deal with the electroweak symmetry breaking and fermion mass generation in a theoretically more elegant way, have been developed, *e.g.* supersymmetric and technicolor models. These models predict the existence of new particles at energies above the reach of past experiments. In order to shed light on these questions it is important to study interactions at the highest possible energies.

The Fermilab Tevatron is the highest energy collider in the world and therefore provides the most powerful tool available to probe the validity of the Standard Model. Precise measurements of the properties of the top (t) quark and the W boson constrain the Standard Model. Deviations may indicate the breakdown of the theory. Large data samples significantly increase the potential to discover new particles, like the Higgs boson or the particles predicted by supersymmetric models.

During Run I, from 1992 to 1996, the Tevatron delivered $p\bar{p}$ collisions at 1.8 TeV to two experiments, DØ and CDF, which both accumulated about 0.1 fb^{-1} of data. The Tevatron is presently undergoing an upgrade which will increase its energy to 2 TeV and the luminosity by an order of magnitude by the year 2000. During Run II, scheduled to begin in April 2000, the Tevatron will deliver an integrated luminosity of 2–4 fb^{-1}. Future upgrades (TeV33) are envisioned to further increase the integrated luminosity by 20 fb^{-1} before the startup of the LHC.

The inelastic $p\bar{p}$ cross section is large, about 50 mb at the Tevatron. At Run II luminosities this results in an interaction rate of 10 MHz. Most of these interactions are without interest to the physics program. For example W bosons are produced only at a rate of about 5 Hz and top quark pairs at a rate of 3 per hour. Sophisticated pattern recognition algorithms are required to separate the interesting events from the overwhelming background. In many rare, but interesting processes, *e.g.* top quark decays, decays of the Higgs boson or other predicted but unobserved particles, b quarks are produced. The detection of b quarks is thus an important tool for tagging interesting events. The b quarks produced in $p\bar{p}$ collisions live long enough (about 1.5 ps) to travel several millimeters. The secondary vertex from the b quark decay is therefore separated from the primary $p\bar{p}$ interaction vertex. With precise charged particle detectors these separated vertices can be detected. Their presence provides a means of tagging events in which b quarks were produced.

2.2 Studies of the Top Quark

For the reconstruction of top quark decays the tagging of b quark jets will play a central role. Top quarks are predominantly produced as $t\bar{t}$ pairs and decay via $t \rightarrow Wb$, so that every $t\bar{t}$ decay contains a b and a $\bar{b}$ quark. We can use b-tagging to reject backgrounds to $t\bar{t}$ decays. In a 4 fb^{-1} data set we expect to identify several thousand $t\bar{t} \rightarrow \ell\nu$+jets events with at least one b tag. This sample of $t\bar{t}$ decays will allow us to study the production dynamics and measure the decay branching ratios of the top quark. If the top quark, for example, couples to an unobserved heavy state, as some theories suggest, it will manifest itself in the shape of the $t\bar{t}$ mass distribution.

The top quark mass is of particular interest. It plays an important role in radiative corrections to many processes. The Standard Model prediction for the mass of the W boson includes radiative corrections which depend on the masses of the top quark and the Higgs boson. Figure 1 shows the predicted W boson mass as a function of the top quark mass for several assumed values of the Higgs boson mass. The data point indicates the world average measurements of the W boson mass [2, 3] and the top quark mass [4, 5]. The measurement of the W boson mass is one of the most sensitive tests of the Standard Model since its value follows directly from the structure of the model. Assuming the Standard Model as correct, the measurements can be translated into a constraint on the mass of the unobserved Higgs boson. The top quark mass is also interesting because of its sheer size, 174 GeV. The top quark is as massive as a gold atom. By far the heaviest known fundamental particle, the top quark may well hold the key to understanding electroweak symmetry breaking and the generation of the fermion masses.

The most precise top quark mass measurements from Run I use the $t\bar{t}$ decay channel in which one top decays to $q\bar{q}b$ and the other to $e\nu b$ or $\mu\nu b$, the "lepton+jets" channel. The precision of the DØ measurement (about 7 GeV) is limited by the number of events, the jet energy scale calibration, and the combinatorics due to the large jet multiplicity. Each of the four final-state quarks fragments into a jet of hadrons. If light quark and b jets cannot be distinguished, there are 12 different ways to assign the four jets to the decay of the t and $\bar{t}$ quarks. If both b jets are tagged this reduces to a mere two-fold ambiguity, significantly reducing the associated systematic uncertainties. The jet energy scale uncertainties can be reduced by reconstructing W or Z boson decays to jets. The most promising channel is the decay $Z \rightarrow b\bar{b}$, in which b tagging can provide rejection of the dominant two-jet background. This rejection is already needed on-line, when the data are recorded, in order to keep the event rate low enough for the experiment to be able to write all events to a storage medium. With these additional handles on systematic uncertainties we should be able to measure the top quark mass to about 3 GeV precision in Run II.

2.3 Search for the Higgs Boson and Phenomena beyond the Standard Model

With a large data set (> 5 fb^{-1}) a signal for $H \rightarrow b\bar{b}$, produced in association with a W boson decaying to $e\nu$ or $\mu\nu$, can be observed at the Tevatron for $80 < m_H < 120$ GeV [6]. This mass window is extremely interesting from experimental and theoretical points of view. It extends the accessible Higgs mass range beyond the reach of LEP before LHC starts running. A Higgs boson in this mass range is difficult to detect at the LHC. Fits to the present electroweak data, assuming the completeness of the minimal Standard Model, give $m_H = 115^{+116}_{-66}$ GeV [3]. In supersymmetric extensions to the Standard Model the lightest Higgs boson is expected to have a mass below about 130 GeV. It is therefore imperative that the window $80 < m_H < 130$ GeV be covered reliably. This will be possible first at the Tevatron. While the integrated luminosity expected

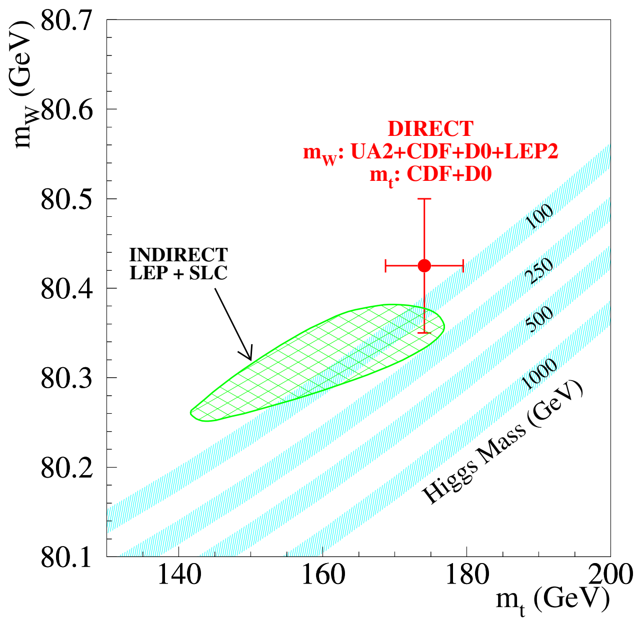

Figure 1: Predicted W boson mass versus the top quark mass (bands).

for Run II falls somewhat short of the amount required to see a Higgs signal, it is close enough that a search is worthwhile. The inclusion of additional channels, like Higgs bosons produced along with hadronically decaying W or Z bosons, will increase the sensitivity beyond the study in reference [6]. In order to exploit these channels we have to be able to store all events with $q\bar{q}b\bar{b}$ final states for analysis. This, like the $Z \to b\bar{b}$ signal, requires on-line rejection of backgrounds without b quarks.

A study of $Wb\bar{b}$ final states may be interesting for other reasons as well. For example, it has been suggested [7] that the decay $\rho_T \to W\pi_T(\to b\bar{b})$ could provide evidence for technicolor. Technicolor theories provide an alternative mechanism for electroweak symmetry breaking that does not involve a Higgs boson. They predict additional pion-like particles ("technipions": π_T) which may decay into b quarks and should be detectable at the Tevatron.

2.4 Bottom Quark Physics

The study of the b quark itself is a topic of significant interest because of the potential for significant CP violation in the B^0–$\overline{B}^0$ system. CP is the symmetry that relates particles to antiparticles. A CP violating physical process proceeds differently if particles and antiparticles are interchanged. The study of the b quark will be key to understanding the mysterious matter-antimatter asymmetry in the universe, which is responsible for our very existence. CP violation has only been observed in the neutral K meson system and its origin is not explained by the Standard Model. Given the ability to write enough b decay data to storage, DØ may be able to detect CP violation in the B^0–$\overline{B}^0$ system with the Run II data.

3 The DØ Trigger System

3.1 General

The proton and anti-proton beams cross in the center of the DØ detector every 132 ns, resulting in a crossing frequency of 7 MHz. On the other hand the data acquisition system can only read out the detector and write the information to tape at a rate of 50 Hz. The function of the trigger system is the recognition of beam crossings during which an interesting event occurred that should be stored for off-line analysis. In order to perform this function efficiently, the DØ experiment will have a three-level trigger system in Run II.

3.2 Level-1

Level-1 is entirely implemented in hardware. A logic network tests for calorimeter towers with energy above preprogrammed thresholds and hit patterns in the muon system and the CFT that are consistent with tracks above preprogrammed transverse momentum thresholds. This test is complete 4.2 μs after the beam crossing. The front ends are buffered by a pipeline 32 levels deep, which makes the Level-1 trigger deadtimeless. If the signals from an event satisfy one of the preprogrammed Level-1 trigger conditions the digitization of the detector signals is initiated and the event is passed on to the second trigger level. The maximum rate at which Level-1 can accept events is 10 kHz. As soon as the digitization is initiated the detector stops accepting new events.

3.3 Level-2

Level-2 consists of dedicated preprocessors, each of which reduces the data from one detector system (calorimeter, muon system, CFT, preshower detectors), and a global Level-2 processor which collects the data from the preprocessors and takes the trigger decision based on the combined information of all detector systems. This decision has to be taken in a very short time in order to limit the deadtime incurred while the Level-2 is processing the event. The mean decision time for Level-2 is 100 μs, resulting in no more than 5% deadtime. Level-2 can accept events and pass them on to the third trigger level at a rate of 1 kHz.

3.4 Level-3

Level-3 consists of a farm of fast hight-level computers and performs a simplified reconstruction of the entire event. Events that satisfy the desired characteristics can be written to a storage medium at a sustained rate of up to 50 Hz.

4 The DØ Silicon Microstrip Tracker

4.1 Overview

Figure 2 shows a view of the SMT which consists of cylindrical barrel sections and disks between the barrel sections and at the two ends of the detector. The barrel sections provide precise measurement of tracks in the central region. The luminous region of the Tevatron has an rms length of 25 cm, which requires the barrel section to be long (72 cm). The disks extend the acceptance of the detector to forward tracks.

4.2 Barrel Detectors

4.2.1 Geometric Layout

The barrel sections consist of rectangular silicon detectors arranged in four layers. Figure 3 shows the arrangement of the detectors in the plane transverse to the beam direction. All detectors have axial strips (parallel to the beam axis).

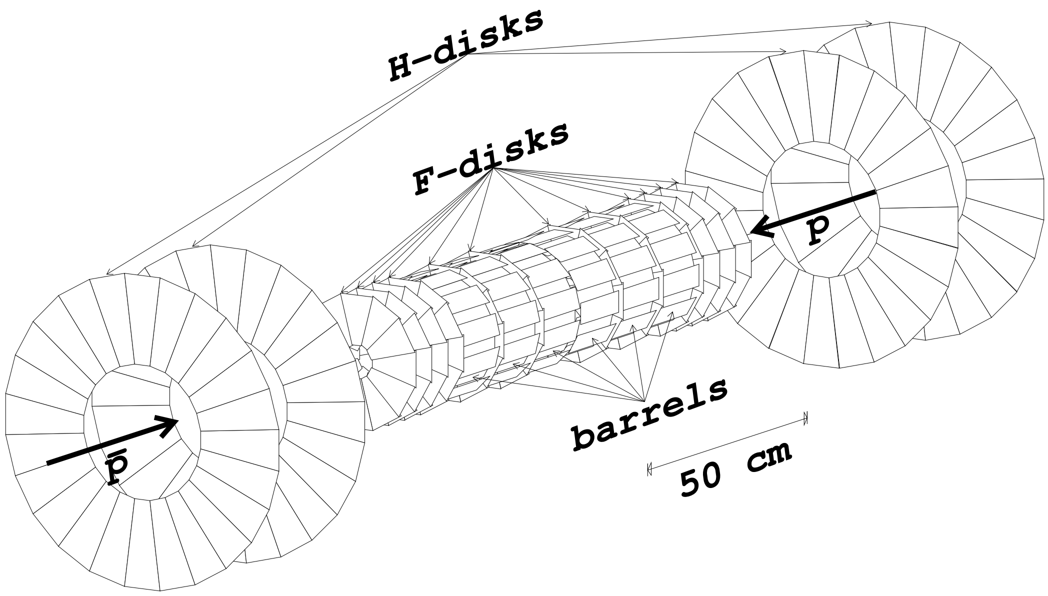

Figure 2: View of silicon microstrip tracker.

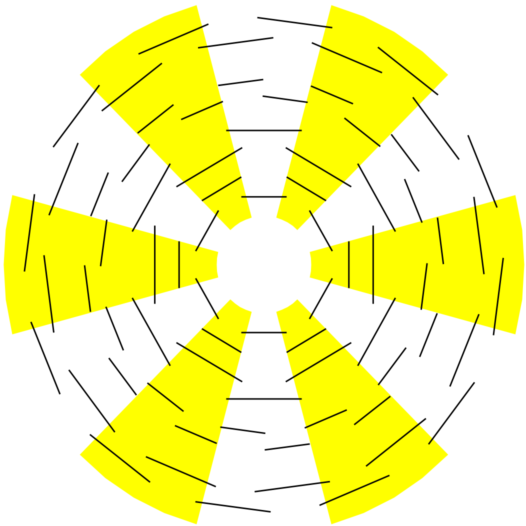

Figure 3: Arrangement of barrel detectors in transverse plane.

4.2.2 Sector Boundaries in SMT

Figure 4 shows three 30° sectors in the silicon detector. The lines indicate the active areas of the ladders and the shaded regions the intervals in azimuth for which tracks with infinite momentum, coming from the origin, are contained in one 30° sector. There is only a tiny interval (1/4°), for which tracks hit two sectors. The acceptance loss from tracks that cross sectors depends in general on their p_T and impact parameter b. Figure 5 plots the acceptance loss versus curvature ($=1/p_T$) of the track for several impact parameter values. The plot of acceptance loss reflects the slight asymmetry of the detector between tracks with positive and negative curvatures and impact parameters. For $|p_T| > 1.5$ GeV the acceptance loss is never more than 1.2% for $b = 0$. For $b = 1$ mm it is on average 1.0% and for $b = 2$ mm it is 2.2%. This losss in acceptance is driven by the small overlap between the active regions of the ladders in layer 3. If tracks with hits in only the other three layers (1, 2, and 4) are accepted, there is no acceptance loss.

We can thus safely ignore tracks that cross between sectors and divide the SMT into 12 independent sectors, each covering 30° in azimuth as shown by the shaded regions in figure 3. The detectors in adjacent sectors overlap so that almost all tracks are contained in a single sector. All but 2% of tracks with $p_T > 1.5$ GeV and impact parameter $|b| < 2$ mm hit detectors in every layer that belong to a single sector. In the STTpp we therefore ignore tracks that cross sectors, which substantially simplifies the design. The STTpp can then be broken down into 12 sectors that do not communicate with each other.

4.2.3 Tolerances

4.2.4 Resolution

4.3 Detector Readout System

4.3.1 Readout Electronics

4.3.2 Pedestals

4.3.3 Signal to Noise Performance

4.3.4 Dead Channels

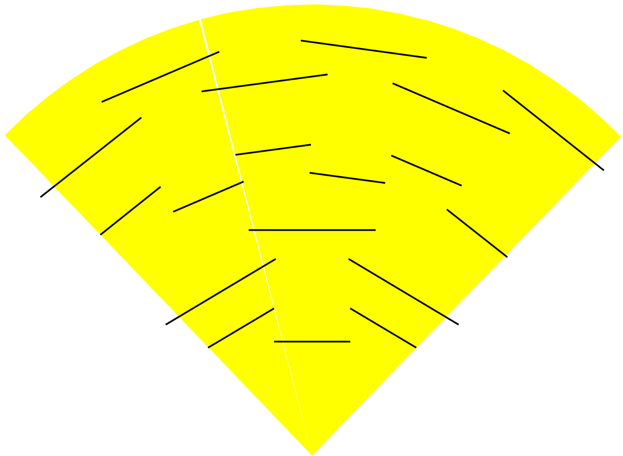

Figure 4: Three sectors in the silicon detector.

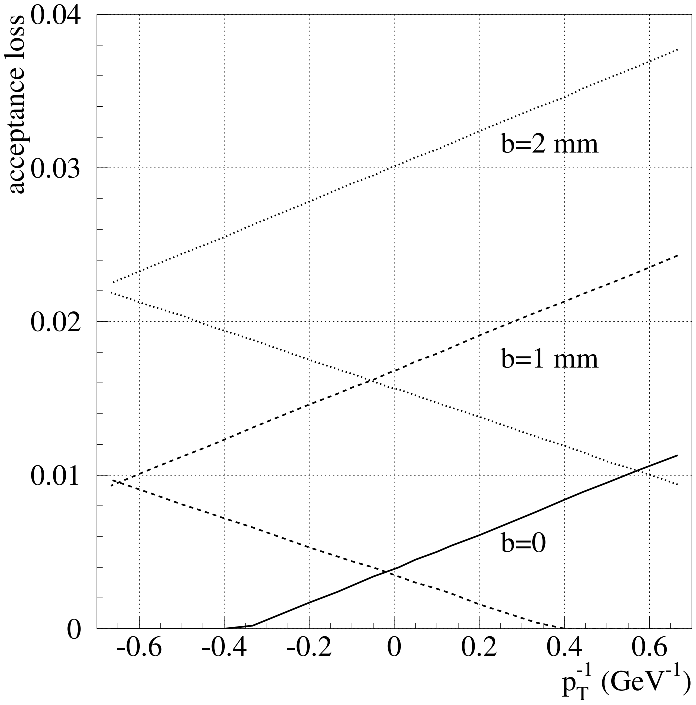

Figure 5: Acceptance loss vs curvature.

5 Conceptual Design

This section describes the conceptual design of the STTpp based on references [8]. The STTpp consists of fast digital electronics that process input data from two detector components: the CFT and the SMT. The CFT Level-1 trigger communicates a list of track candidates to the STTpp. Each track candidate defines a road in the SMT. The STTpp receives all data from the SMT barrel detector portcards. It forms hit clusters from the raw data and filters these clusters, retaining only the clusters that fall into one of the roads defined by the CFT data. Of all the hit clusters in a road, it finds the hits most consistent with the corresponding CFT track and fits a trajectory to the selected SMT hits and the CFT track. Then it communicates a list of fitted track parameters to the Level-2 CFT preprocessor (CFTpp), which merges the information with the Level-1 CFT information and formats the data for transmission to the global Level-2 processor.

As discussed in section 4, the SMT detectors can be divided into 12 sectors, each covering 30° in azimuth. In the STTpp we ignore tracks that cross sectors, which substantially simplifies the design. The STTpp can then be broken down into 12 subunits, each covering one sector, that do not communicate with each other. We have to custom-design the electronics for these subunits in order to optimize them for the required speed. Two such subunits share a VME crate. There are six such crates. In each crate there is one fiber road card, which receives the information from the CFT Level-1 trigger and broadcasts it across the backplane to up to 12 trigger cards. The trigger cards convert the CFT information into roads in the SMT. They receive the SMT data from the portcards, find hit clusters and sort the clusters by roads. The data processing on the trigger cards takes place in field programmable gate arrays (FPGA). The track fitting is performed by digital signal processors (DSP), located on one or more processor cards.

The electronics in the CFTpp, consisting of processors and communication links are copies of components that are already being designed for other parts of the DØ trigger system. In this way the STTpp will fit seamlessly into the existing Level-2 design. Figure 6 shows the crate structure of the STTpp.

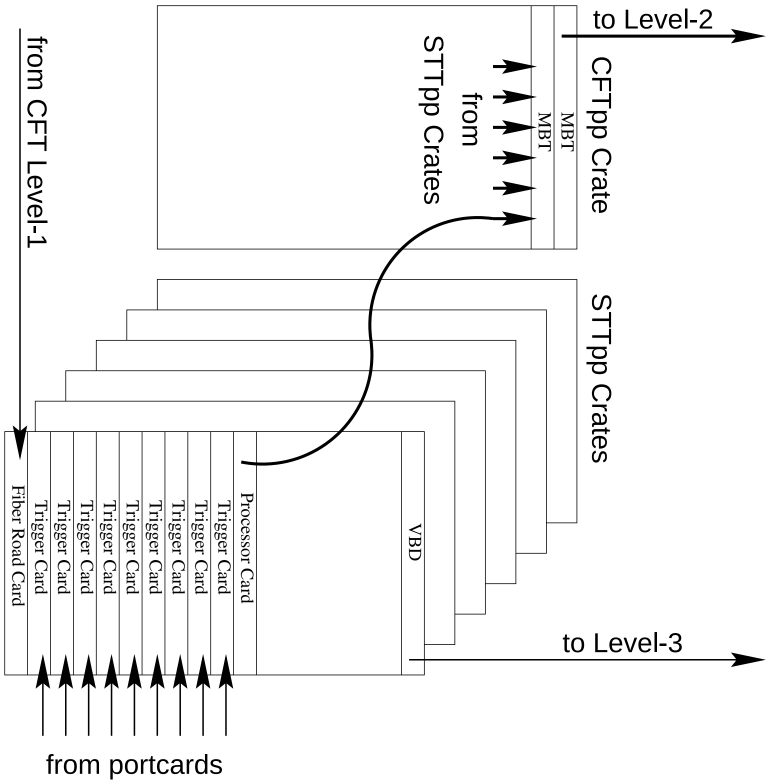

Figure 6: Crate structure of STTpp.

6 Fiber Road Card

6.1 Overview

The fiber road card serves three functions in the overall STTpp design.

- It receives roads from the CFT Level-1 trigger.
- It sends the roads out over a special purpose bus to the trigger cards.
- It is used as the controller and buffer manager for the STTpp crate.

One of these cards will be necessary for each STTpp VME crate, corresponding to one STTpp sector. A block diagram of the fiber road card is given in Figure 7.

6.2 Input from CFT Level-1 Trigger

The CFT Level-1 trigger transmits up to 32 CFT track candidates over a single fiberoptic cable to each STTpp sector 60° wide. For each track candidate it sends the identities of the fibers hit in the inner and outermost CFT layers. Tracks that pass through a given STTpp sector can hit any one of 773 fibers in the outermost CFT layer, spanning a region of 60°+19° in azimuth [9]. Tracks with $p_T > 1.5$ GeV that pass through a given outer fiber can hit any one of 28 fibers in the inner CFT layer. We thus need at least 15 bits to completely specify a CFT track candidate, 10 to identify the outer fiber and another 5 for the inner fiber.

The exact scheme to map the 80 azimuthal CFT sectors onto the six 60° STTpp sectors, that looses as few as possible tracks which cross sector boundaries, is currently under study.

The CFT Level-1 track lists will be obtained by passively splitting the fibers that transmit the CFT track lists from the CFT Level-1 concentrator cards to the CFTpp. The track lists will have one 16-bit word for each track.

A fiber optic receiver will be necessary to convert the CFT data to electrical signals. The data will then be put into a FIFO which decouples the data-sending clock from the local crate clock.

6.3 Buffer and Crate Control Functions

In addition to dealing with the standard DØ crate controls sent via the Serial Command Link, the fiber road card must also control buffering and readout of the data from the processor cards to Level-3 via VBD.

6.4 VME Interface

The ability to access data in the card over VME at various stages in the processing for diagnostic purposes is forseen.

7 Trigger Card

7.1 Overview

The trigger card has to fulfill the following tasks:

- receive CFT data from the fiber road card over the backplane and convert it to roads in SMT,
- receive SMT data from the portcards over optical fibers and find hit clusters,
- select SMT hit clusters that are inside the roads defined by CFT data,
- buffer SMT data,
- transmit the hit clusters in the roads to the processor card(s),
- buffer hit clusters for readout to Level-3 over VME bus.

Figure 8 shows a block diagram of the trigger card.

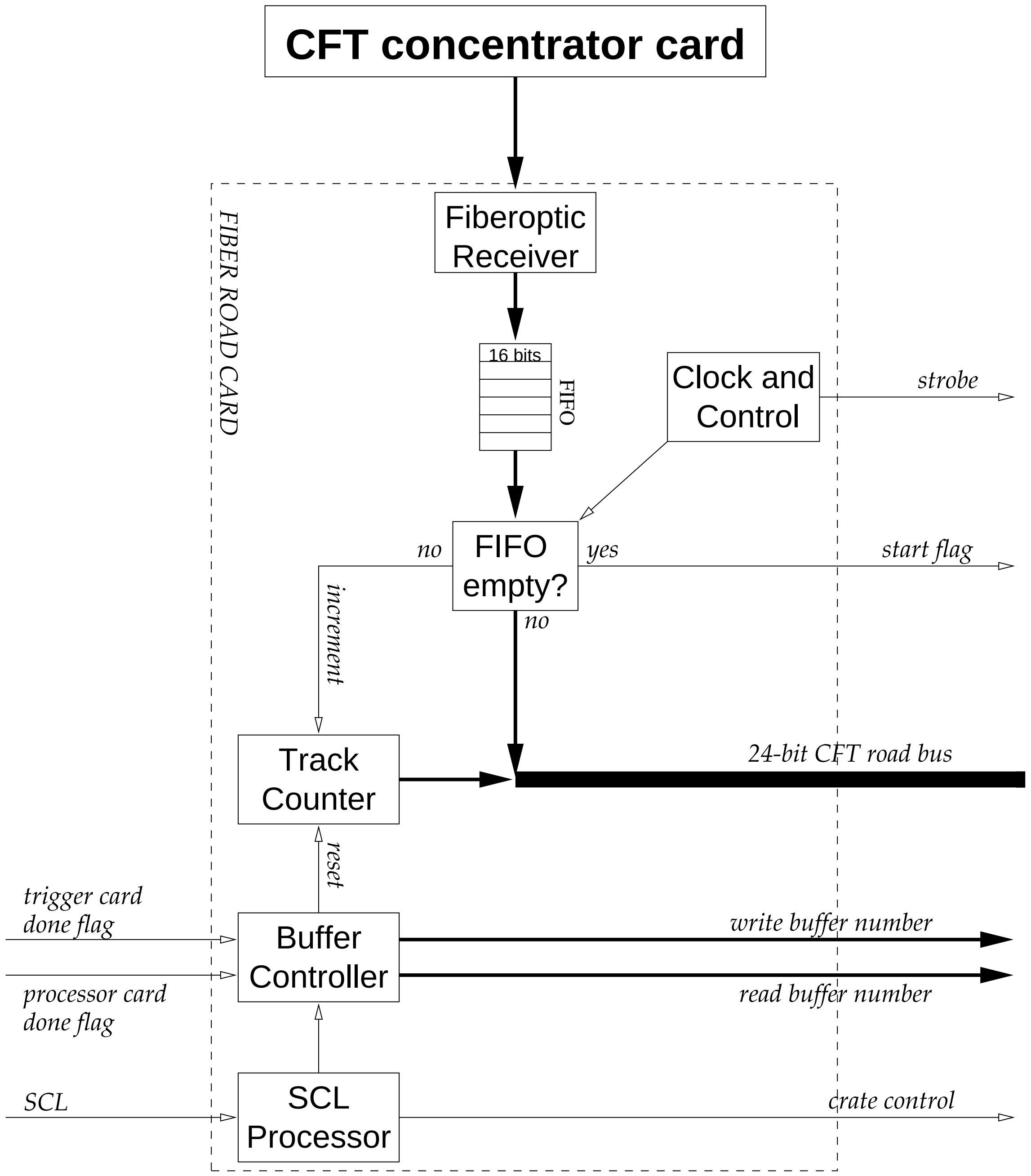

Figure 7: Block diagram of fiber road card.

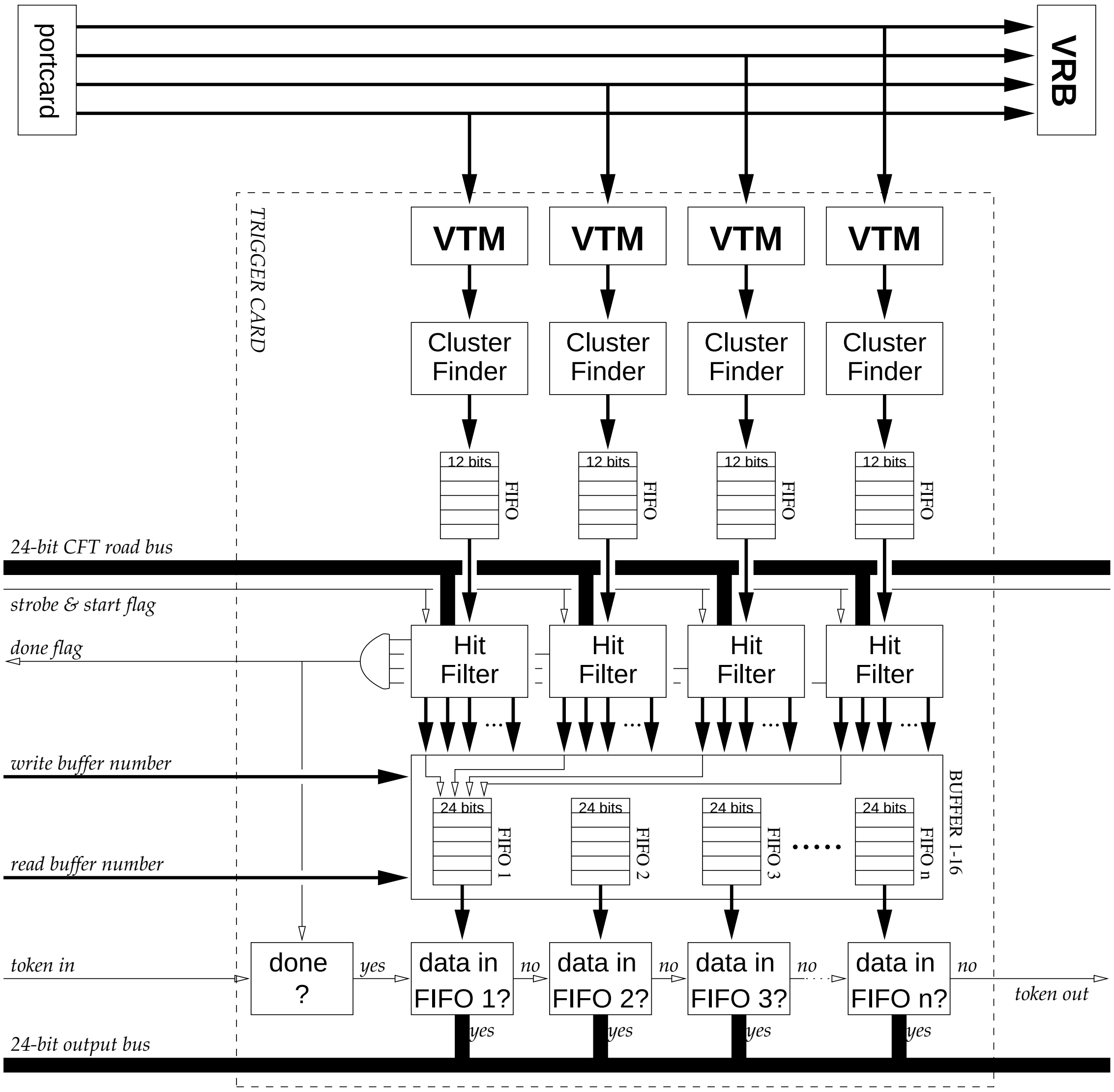

Figure 8: Block diagram of trigger card.

7.2 Input from SMT Front-Ends

Optical fibers transmit the SMT data from the portcards on the detector platform to the VRBs [10] in the counting house. The data transmitted over a fiber consist of alternating 8-bit address and signal words. These words are transmitted at a rate of 53 MHz. On every fiber an end-of-record code indicates that all data for an event have been sent. Passive splitters in the fibers create a data path into the STTpp, where the fibers plug into the same VTM receiver module [11] as used for the VRBs. The data are loaded into a FIFO as they are received.

Each barrel detector is read out over a single HDI. Each portcard accepts data from eight HDIs and outputs it over four optical fibers such that two HDIs map into one fiber. The four fibers from one portcard plug into the same VTM at the VRB end and at the STTpp end. The mapping of the barrel detectors into portcards must therefore arrange the SMT data in the way that the STTpp requires it.

The STTpp has to provide seamless coverage of all barrel segments that belong to one SMT sector, because many tracks cross from one barrel segment into another. Very few tracks are not contained in a single SMT sector (section 5) so that each sector can be treated as an independent unit. This means that the data from all barrel segments of any one SMT sector must be routed to the same STTpp crate. For a six-barrel configuration there are 36 detectors per sector which populate 4.5 trigger cards. The system must be extensible to an eight-barrel configuration with 48 detectors per sector which populate 6 trigger cards. We can therefore accomodate two SMT sectors in each STTpp crate. Table 1 shows how the detectors from the two sectors map onto the trigger cards.

Table 1: Map of barrel detectors to trigger cards for one STTpp crate (6 barrel segments). The letters A–D label the input fiber within a trigger card.

barrel segment	1	2	3	4	5	6
sector 1 layer 4b	1D	1B	2B	3B	4B	4D
sector 1 layer 4a	1C	2A	2C	3C	3A	4C
sector 1 layer 3b	1C	2A	2C	3C	3A	4C
sector 1 layer 3a	1D	1B	2B	3B	4B	4D
sector 1 layer 2b	5A	1A	2D	3D	4A	5B
sector 1 layer 1a	5A	1A	2D	3D	4A	5B
sector 2 layer 4b	6D	6B	7B	8B	9B	9D
sector 2 layer 4a	6C	7A	7C	8C	8A	9C
sector 2 layer 3b	6C	7A	7C	8C	8A	9C
sector 2 layer 3a	6D	6B	7B	8B	9B	9D
sector 2 layer 2a	5C	6A	7D	8D	9A	5D
sector 2 layer 1b	5C	6A	7D	8D	9A	5D

Mapped in this way each fiber carries data from one 6-chip and one 9-chip HDI for the inner four barrels and one 3-chip and one 9-chip HDI for the outer two barrels.

7.3 Hit Finding

An FPGA processes the data further. It corrects for pedestals and gains and supresses noisy strips based on a lookup table. It supresses all strips with pulseheights below a threshold. It has to be determined whether these operations need to be performed on a strip-to-strip, chip-to-chip, or global scale. The FPGA then scans the corrected pulseheights for local maxima, *i.e.* strips with larger pulseheights than their two neighbors. Each such local maximum forms the seed for a hit cluster. The position of the hit cluster is determined from a pulseheight averaged center of gravity of up to five adjacent strips, centered on a seed strip, using the following algorithm. Hit clusters cannot extend across two detectors nor can they extend over both sides of a double sided detector.

Let p_i be the 8-bit pulseheight measurement in strip i. The strip index i is defined relative to the strip with the local maximum, which has $i = 0$. The 7-bit strip number is x_i. Compute

$$\begin{aligned} n &= 8 \times \sum_{i=-2}^{2} p_i i \text{ and} && (1) \\ m &= \sum_{i=-2}^{2} p_i . && (2) \end{aligned}$$

Then set the offset Δx using the following comparisons

$$\begin{array}{rcll} & n & > 5m & \Rightarrow \Delta x = 3/4, \\ 5m > & n & > 3m & \Rightarrow \Delta x = 1/2, \\ 3m > & n & > m & \Rightarrow \Delta x = 1/4, \\ m > & n & > -m & \Rightarrow \Delta x = 0, \\ -m > & n & > -3m & \Rightarrow \Delta x = -1/4, \\ -3m > & n & > -5m & \Rightarrow \Delta x = -1/2, \\ -5m > & n & & \Rightarrow \Delta x = -3/4. \end{array} \tag{3}$$

The cluster centroid is then given to a precision of 1/4 strip pitch by

$$\overline{x} = x_0 + \Delta x. \tag{4}$$

This algorithm avoids floating point divisions which are difficult to implement in an FPGA. The cluster outputs 13 bits: the seven bits used to identify the strip number plus an additional two added by the cluster finding algorithm, three bits to label the eight chips that read out the axial strips or the seven chips that read out the stereo strips, and one bit to indicate whether the cluster is on the axial or the stereo side (see figure 9).

13	12	11	10	9	8	7	6	5	4	3	2	1
axial stereo	chip			strip								

Figure 9: Bit allocation in output of cluster finder.

Based on bit 13 the clusters are stored in two FIFOs. Only clusters from the axial strips participate in the track finding. The axial strips are read out by the first (last?) three chips for layers 1,2,5, and 6, and by the first (last?) five chips for layers 3,4,7, and 8. The clusters from the axial strips are stored in one FIFO, the clusters from the stereo strips are stored in the other FIFO. All clusters are buffered for readout to Level-3 over the VME bus (?).

7.4 Input from Fiber Road Card

The data from the CFT Level-1 trigger are transmitted from the fiber road card over a bus on the backplane. For every CFT track 24 bits have to be transmitted: 16 bits to identify the inner- and outermost hit fibers and an 8-bit track number. The bus requires 24 signal lines and one strobe line. It runs at 25 ns per cycle. According to Monte Carlo simulations [12], we expect on average less than two tracks per STTpp crate. The maximum number of tracks expected is 25, the maximum number allowed per STTpp crate is 32. It therefore takes 50–800 ns to transmit the CFT data across the backplane to the trigger cards.

7.5 Conversion of CFT Tracks to SMT Roads

The 16 lines that identify the CFT track address a lookup memory. The output of the lookup memory is a range of strip numbers, consisting of 10 bits each (corresponding to bits 3–12 in figure 9) for the first and last strips. If a track does not go through a detector the corresponding lookup memory will return zero for both strip numbers. Since every input fiber carries the data from two barrel detectors, up to eight strip number ranges have to be determined for every CFT track. Tracks that go through an SMT sector of 60° in azimuth can hit any one of 773 outer fibers. For each outer fiber there are 28 inner fibers that can be hit. So there are 21644 possible CFT tracks for a 60° SMT sector. We therefore need either one lookup memory of 433 kB with 16 address lines and 160 output lines or eight lookup memories of 54 kB each with 16 address lines and 20 output lines. The coordinate conversion can then be performed "on the fly" just like the cluster finding and no time is lost.

7.5.1 Road Definition

The fiber hits in the inner and outer CFT layers define a family of tracks, which all go through these fibers. The envelope of all these tracks defines the road in the SMT. For radial positions smaller than that of the inner hit fiber, the tracks with the smallest and largest curvature define this envelope. Figure 10 shows the road in the SMT spanned by two hit fibers in the CFT. The solid lines indicate the road for tracks that pass through the origin. If we also want to accept all hits from tracks with finite impact parameters b the road widens as indicated by the dashed lines.

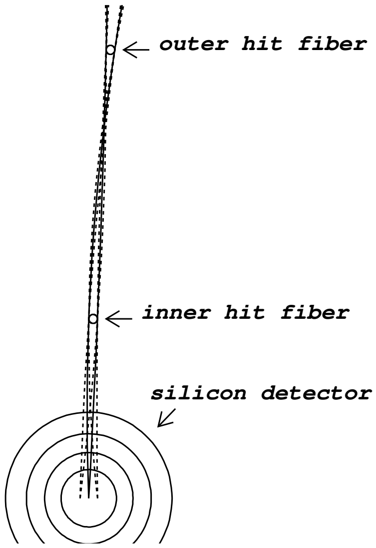

Figure 10: Road definition in the silicon detector.

Figure 11 shows the width of the road in azimuth around the track, defined by the centers of the inner and outer fibers and the origin, versus radius. The black region is the road defined by the envelope of all tracks passing through the origin, the dark shaded region is the road defined by tracks with impact parameter $|b| < 1$ mm, and the light shaded region for $|b| < 2$ mm. Figure 12 shows the road widths in arclength. Table 2 lists the road widths at the mean radii of the silicon layers. In the table "a" and "b" refer to the two overlapping sublayers that make up one silicon layer.

7.6 Hit Filter

As the cluster data appear at the output of the cluster-position FIFO, they are processed by a second FPGA. This FPGA simultaneously compares the cluster positions to the upper and lower limits of the all address ranges for all roads. Bits 3–12 in figure 9 are compared to the road limits. Since there can be up to 32 roads

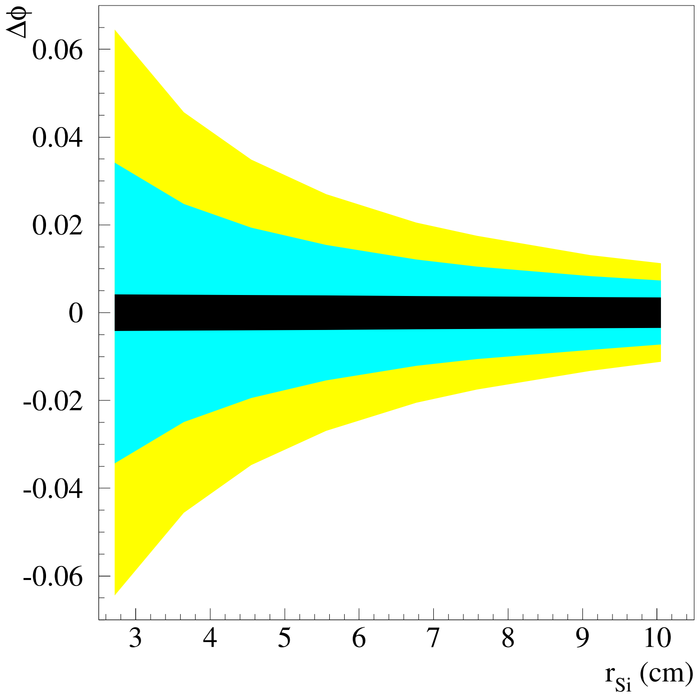

Figure 11: Width of silicon roads versus radius.

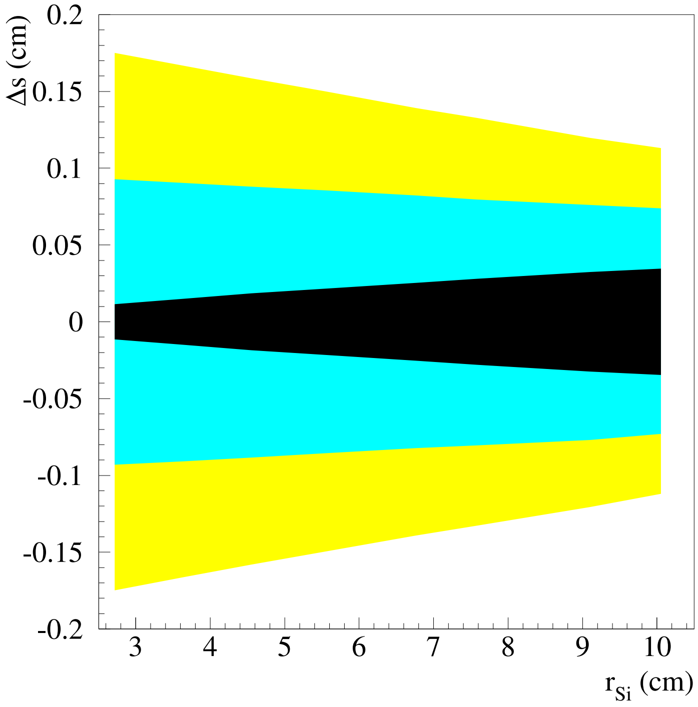

Figure 12: Width of silicon roads versus radius

Table 2: Road sizes in silicon detector.

layer	radius	road for $b = 0$		road for $\|b\| < 1$ mm		road for $\|b\| < 2$ mm	
		mrad	mm	mrad	mm	mrad	mm
1a	2.72 cm	$\pm$ 4.2	$\pm$0.12	$\pm$34.3	$\pm$0.93	$\pm$64.4	$\pm$1.75
1b	3.65 cm	$\pm$ 4.1	$\pm$0.15	$\pm$24.8	$\pm$0.91	$\pm$45.7	$\pm$1.66
2a	4.55 cm	$\pm$ 4.0	$\pm$0.18	$\pm$19.4	$\pm$0.88	$\pm$34.8	$\pm$1.58
2b	5.55 cm	$\pm$ 3.9	$\pm$0.22	$\pm$15.4	$\pm$0.86	$\pm$27.0	$\pm$1.50
3a	6.77 cm	$\pm$ 3.8	$\pm$0.26	$\pm$12.1	$\pm$0.82	$\pm$20.6	$\pm$1.39
3b	7.58 cm	$\pm$ 3.7	$\pm$0.28	$\pm$10.5	$\pm$0.80	$\pm$17.5	$\pm$1.33
4a	9.10 cm	$\pm$ 3.6	$\pm$0.32	$\pm$ 8.4	$\pm$0.76	$\pm$13.2	$\pm$1.20
4b	10.05 cm	$\pm$ 3.4	$\pm$0.35	$\pm$ 7.3	$\pm$0.73	$\pm$11.2	$\pm$1.13

and two address ranges per road per fiber, a total of 128 10-bit comparators are required. Figure 13 shows a block diagram of the hit filter FPGA.

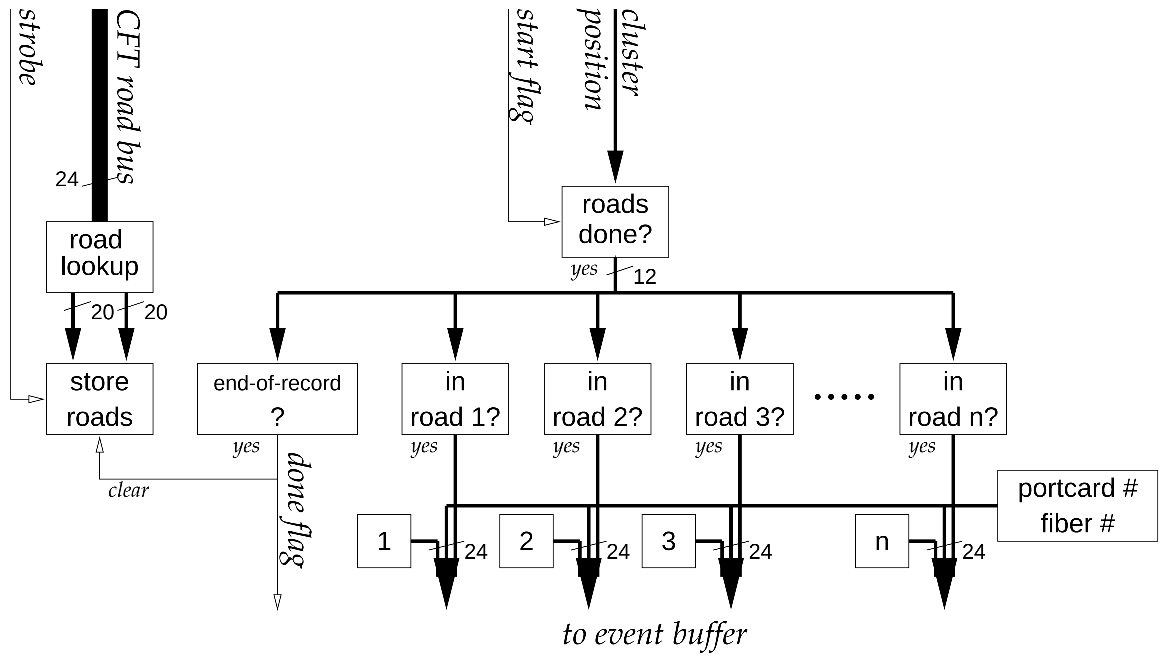

Figure 13: Block diagram of hit filter FPGA.

7.7 Event Buffering

The output of the hit finder FPGA is stored in a buffer 16 events deep. There is a FIFO for every road that stores all (12-bit) cluster positions that are inside an address range for this road. An end-of-record code is passed through the cluster finder unaltered. When it appears at the hit filter, it sets a "done" flag to indicate that all data from the fiber has been processed through the cluster finder and the hit filter and has been loaded into a buffer. The logical "and" of the flags from the four hit filters on a trigger card indicates that the trigger card has finished loading an event. The logical "and" of the "done" flags from all trigger cards is formed on the backplane (?) and signals to the fiber road card that an event has been completely loaded into the buffer. The fiber road card then updates the buffer pointer for the next event to point to an empty buffer. At this time the STTpp is ready to accept another event.

7.8 Timing

It takes about 2.4 μs after a Level-1 trigger accept for the CFT data to arrive at the fiber road cards. Bus transfer to the trigger cards takes another 0.8 μs. That means 3.2 μs after a Level-1 trigger accept the road addresses are latched on the trigger cards.

Digitizing the SVX-II data takes a little under 3 μs and there is an additional small latency of about 100 ns from the cluster finding. Thus the SMT data arrive in the FIFO at the output of the cluster finder roughly at the time the roads are set up.

Every fiber carries data from about 2000 SMT channels. Monte Carlo simulations predict an average of 40 hits per fiber, corresponding to an occupancy of 2% [12] and a maximum of 180 hits per fiber. We add 1% occupancy from noise (see analysis of test beam data [13]). We thus have 60–200 channels to read out. At 53 MHz this will take 2.3–7.6 μs. Only half the hits (from the axial strips) form clusters and the reduction from number of hits to number of clusters should be about a factor two. Thus there will be 15–50 clusters which have to be compared to the strip ranges that define the roads. The comparisons for all roads are performed in parallel. For each road there is a FIFO and a cluster that is in the strip range for a road is stored in the corresponding FIFO. This should not take longer than 100 ns (?) per cluster, *i.e.* 1.5–5.0 μs. This is less time than required for the readout of the SMT. The cluster filtering can therefore effectively be performed without latency.

All hit clusters in the roads for an event are loaded into a buffer about 5. 5–10.8 μs after the Level-1 trigger accepted the event, but not later than completion of the readout of the SMT data. At this time the STTpp is ready to accept another event from Level-1.

7.9 Transfer of Hits to Processor Card

For each hit 26 bits need to be transferred: 7+2 bits for the fractional strip number, 3 bits to identify the HDI/chip within a fiber, 2 bits to specify the fiber within a portcard, 4 bits to identify the portcard, and a 8-bit track number (figure 14). The lower 12 bits are the same as at the output of the cluster finder (figure 9).

26 ... 19	18 ... 15	14 13	12 ... 10	9 1
track	portcard	fiber	HDI/chip	strip

Figure 14: Bit allocation in output to processor card.

Hit transfer to the processor card is initiated by the presence of data in the event buffer. The first trigger card transmits the lists for all roads over a 26-bit bus to a processor card. For each road the first record consists of the CFT data and the road number. These are the 24 bits sent from the fiber road card to the trigger cards. Then all hits in the road are transferred. This sequence is repeated for all roads. When the card has finished transmitting the list it passes a token to the next trigger card and so on until all trigger cards that serve the same azimuthal sector in the SMT have transmitted their data to the processor card.

There are on average 42 and at most about 500 clusters in all roads of an STTpp crate [12]. If the bus operates at 25 ns/cycle, the transmission of these clusters and the CFT data to the processor card will take 1.1–13.3 μs.

7.10 VME Interface

7.10.1 Parameter Download

The VME interface has to support the download of the lookup tables for the conversion of CFT coordinates to address ranges in the SMT for the roads, and

7.10.2 Diagnostics

In mark-and-pass mode all cluster and road information is transmitted to Level-3 over the VME bus and a VRB.

8 Processor Card

8.1 Hit Conversion and Track Fitting

The match between CFT trigger tracks and clustered silicon information is made in the STC. The next stage of processing – silicon coordinate conversion and track fitting – is described in this section. It contains an overview of the fitting system followed by a description of each of the three functional blocks.

Figure ?? shows a block diagram of the track–fitting card. There are three basic functional blocks. They are:

- an **input stage** which receives combined SVT cluster and CFT track information and provides conversion from cluster strip coordinates to global coordinates needed for track fitting,
- a **fitting stage** consisting of an array of dedicated processors to perform the track fit dedicating one processor to each track, and
- an **output stage** which communicates the fit results to the rest of the trigger system.

Because many details remain to be optimized, the design presented here should be seen as an initial, but workable, solution. The most important pieces of the track fitting process have been considered, and the resulting design is believed to be sufficient. In particular, the processor design presented here is based on Texas Instruments TI320C62 digitial signal processors(DSP). These may be replaced in the final design by programmable logic devices, most probably the Altera 10k100 device. However, this would be driven by price and performance considerations, not, we believe, by problems intrinsic to the DSP–based design presented here.

8.1.1 Input Stage

The input stage receives the matched SVT cluster and CFT track information from the silicon trigger card over a dedicated bus. The bus provides a uni–directional, data–driven broadcast from the trigger cards to the track fitting card. When the cluster finding and road matching in the STC is complete, each cluster and the ID of the associated CFT track are driven onto the bus. Each transfer is 30 bits wide, and the cycle time is 25 ns. The cluster information includes the barrel number (3 bits), ladder number (7 bits) and cluster position in strip coordinates (12 bits)[1]. The track ID is taken to be a global serial number uniquely identifying each track for a given event. For a system capable of handling 200 tracks, eight bits are needed for the track ID. In addition to the cluster information, start– and end–of–data transfers will provide synchronization and error checking, and CFT track parameters will be sent as needed by the fitting procedure.

The input stage consists of bus receivers, simple control logic, a table to provide track ID to processor ID translation, and a look up table to provide conversion from hits positions in ladder coordinates (barrel, HDI, SVX chip ID and centroid position) to positions in global (r, ϕ) coordinates.

There will be one processor for each track to be fit, with an upper bound of 32 tracks for each 60^o sector. Each silicon hit will be broadcast with a global track ID indicating which CFT track the hit is associated to. The track ID is a global ID, and a look up table is used to associate a given track to the appropriate processor. A global track ID is input to the look up table, and the ouput is the ID of the processor handling that track. The table is reset at the start of each event and built on–the–fly as data for each track is received. The needs are very modest, 32 8-bit comparators, 32 8-bit registers and minimal control logic.

In addition to the look up table for track ID to processor ID conversion, a second table is used to convert the cluster into global coordinates for the track fit. We have adopted the simplest possible approach to

[1] It is likely that the cluster information will be in the format channel number, SVX chip ID and HDI side. In this case the number of bits can be slightly reduced.

perform this conversion. A look up table is built containing one entry for every possible cluster position in the geometrical sector. This implies between 128k and 256k entries in the look up table for one 30^o sector. The address is formed from the barrel, HDI and SVX chip ID's and the cluster position in strip coordinates. Each entry contains the aligned hit position in (radius,phi) space. If the radius is stored in units of 10 microns and the phi angle granularity is derived from the smallest angular resolution resulting from the 1/4 strip cluster precision, the table needs at most 32 bits for each slot. This corresponds to 14 bits for radius and 18 bits for the angle.[2]

The cost of such a table implemented in 15 ns static RAM will be less than \$70 per 30^o sector. Processor clock cycles are typically 5 ns, considerably faster than the external bus cycle, so converted hits can be copied directly into processor memory.[3]

Although the coordinate conversion table is quite large, the download is not a significant problem because it changes only when the detector alignment changes. Battery backed up RAM could be used to eliminate the need to download after power failures.

8.1.2 Processing Stage

During the track/hit loading phase, the input stage is used to fill processor internal memories with the initial track parameters and positions of associated silicon hits. When this is completed the processors are started, and the tracks are fit using the silicon information. The fit algorithm studied thus far is a linearized fit to the function

$$\phi(r) = b/r + \kappa r + \phi_0$$

This is the same algorithm used in the physics studies. The results from the fit are the three track parameters b, ϕ_0 and curvature and the χ^2/dof. Here b is the vertex-signed impact parameter and ϕ_0 is the angle of the track at the point of closest approach. The algorithm used in most of the physics studies constrains the hits such that the barrels must be in a mechanically sensible order, and if the χ^2 is too large, the track is refit, dropping the hit dominating the χ^2.

The processors will be commercial units, either digital signal processors or programmable logic. The simplest approach would likely be the use of DSP's, but the final choice will be driven by overall cost and performance. At this point in the design, it is important to ensure that the fit can complete within the overall L2 timing constraints. We are conservatively allowing 10 μs for the fit. The time taken has been estimated for one possible processor, the Texas Instruments TMS320C62, by counting instructions needed by the fit. Simulations are underway using silicon Monte Carlo and DSP compiler/simulation packages from TI.

To perform the timing estimates, the fortran code used in the physics studies was reordered to give an apparently minimal number of operations. The number of operations was tallied under the assumption that in DSP processors, additions following multiplications take the same amount of time as multiplications. It was further assumed that the initial track parameters were available in memory, and that hit resolutions were also available as look up. Under these assumptions, the number of operations per track N was found to be

$$N = 27n \text{ multiplies} + (3+n) \text{ divides}$$

in which n is the number of hits on the track.

A simple estimate of the dynamic range needed during the fit gives 35 integer bits, beyond the 32 bit integer capacity of the processor. This implies that floating point formats will be needed[4]. In estimating the fit times, it is assumed that the emulation of 32-bit floating point operations in 32-bit integer format requires 10 instructions for each floating point operation. Roughly, this corresponds to two loads, two shifts, an integer multiply, an integer sum, a shift and a store. This gives eight operations, which is taken as 10 here.

[2] It will be important to insure that the mapping of discrete phi values in the table meshes with the discrete precision of the input clusters. One way to avoid possible mismatches would be to store a ladder ϕ and a cluster $\delta\phi$ in the look up table. The true position would be the sum of the two numbers. This, however, does introduce the need for an addition operation in the conversion.

[3] If synchronizing the loading of the table outputs into the assigned processor is difficult to complete within the bus cycle, the output could be latched for a single cycle. The processor load would then occur during the following cycle. This is essentially creates a two-stage pipeline, and has no real impact on performance.

[4] Other possibilities such as dual range calculations will be studied. If possible, integer execution will be used.

The DSP under consideration is rated as a 1600 MIPS processor, but local tests indicate the processing capacity is 400 MIPS. The external clock frequency is 200 Mhz. The processing time is given as

$$\begin{aligned} T &= (\text{Numberofinstructions})(\text{timeperinstruction})(\text{emulationfactor}) \\ &= (27n + 3 + n) * 2.5 \text{ ns} * 10 \end{aligned}$$

For a fit using four silicon hits, the corresponding processing time is 2.8 μs. Thus, it appears that single–pass fitting can be completed well within the alloted 10 μs. Two–pass fitting in which the worst hit is discarding for tracks with a large χ^2 is also apparently feasible but given the precision of this result two–pass fitting clearly needs more study. Furthermore, this result assumes that only one hit per layer is available. In practice this filtering is best performed after hit conversion because misalignment effects can be considered. The feasibility of performing both hit filtering and a multiple pass fit remains an open question.

8.1.3 Output Stage

The ouput stage is used to transfer the fit results to the global DAQ. The transmitted information consists of the three track parameters and the χ^2 resulting from the fit. In addition, one may want to read out the hit positions for a subset of events. The output will either be to a dedicated local card via a custom bus or directly to a concentrator crate via a serial link.

An upper bound on the data transfer can be derived under the assumption of 32 bits of precision for each of the four quantities from the track fit. At 50 tracks per event, and 16 bytes per track, 800 bytes would be transferred. However, the output can probably be packed to reduce size as follows:

- **Impact Parameter** storing with 1 μm precision and a range from -1.0 cm $\leq b \geq$ 1.0 cm requires 15 bits,
- Momentum storing with 0.1 GeV precision and a range from -100.0 to 100.0 GeV requires 11 bits,
- ϕ_0 storing with 0.025 rad precision requires 8 bits
- χ^2/dof storing with a precision of 0.2 and a range from 0 to 10.0 requires 6 bits.

For a 50 track event, this results in 250 bytes transferred.

8.1.4 Plans

An initial design for the track fitting process has been presented. It includes the functionality needed to convert hits from a silicon–specific coordinate system to the global system and the processing necessary to perform a track fit. Before beginning more detailed design, the fit algorithm timing will be better understood. This, and the resulting processor choice, are the most pressing issues to be resolved.

9 Communication with Global Level-2 Processor

9.1 Track Preprocessor Crate Overview

The results of the track fits performed on the processor card will be sent to MagicBus Transciever cards (MBT) [14] in the CFTpp crate using the Cypress Hot Link serial output bus. One hotlink is envisaged per processor card. With one processor card per crate, the output of the entire STTpp should fit into one MBT in the CFTpp crate.

The exact layout of the Level-2 CFTpp crate is still being decided. However, it is forseen to look similar to the standard DØ Level 2 preprocessor crate. Its main working elements will be

CFT MBTs Receive CFT Level-1 track lists from CFT concentrator crate. Send this information out onto MagicBus.

STT MBT Receive STTpp data from processor cards. Send this information out onto MagicBus.

Alpha Processors Process CFT and SMT information. Transmit results over MagicBus to Level-2 MBTs.

Level-2 MBTs Send processed Level-2 track trigger information to the global Level-2 processor via Cypress Hot Link.

9.2 Input from Processor Cards

Data from the processor cards will contain the information shown in table 3. The exact size of some of the data quantities will depend on the algorithm used for the track fit, and is thus not yet defined. However, each should fit in a 16-bit word. Monte Carlo studies [12] indicate that the average number of tracks with STT information should be around 1 per sector while the maximum possible would be 32. Transfer speed to the MBT will be 160 Mb/s meaning that transfer times for the track parameters would average 0.8 μs and could reach as high as 26 μs.

In addition to the track parameters it is also desirable to have cluster information available at the CFTpp. The implications of this option are discussed in more detail in the next section.

Table 3: Data transmitted to the CFTpp crate per track. Also given is the anticipated size of each quantity per track in bits and where the quantity is originally defined.

Quantity	Bits	Source
CFT track info and number	24	fiber road card
b (signed)	16	processor card
p_T	16	processor card
ϕ_0	16	processor card
χ^2	16	processor card
x position of beam	16	processor card
y position of beam	16	processor card
layers used bit mask	4	processor card
padding (?)	4	
total	**128**	

9.3 Data Processing in the CFTpp and the global Level-2 Processor

Output of the results of the CFTpp processing to the global Level-2 processor will be done over Cypress Hot Links from MBT(s) in the CFTpp crate to an MBT in the global Level-2 processor. Algorithms for combining the STT information with the track trigger are currently under discussion, as are the details of exactly what level of processing will be done with the data in the CFTpp and global Level-2 processor.

10 Communication with Level-3

10.1 Readout Requirements

According to DØ rules all data created by a processor must be read out by Level-3 on every event. The STTpp creates two kinds of data: hit clusters tracks. The raw SMT data are read out through the VRB crates and the CFT Level-1 track candidates are read out from the CFTpp.

10.2 Readout Options

10.2.1 Readout from Trigger Cards and CFTpp

To save unpacking time in Level-3 it may be desirable to read all hit clusters out. These data are only available on the trigger cards, so that in this case the trigger cards have to be read out. Assuming that there are as many clusters on the stereo side of a double sided detector as there are on the axial side, we expect a total of 260 clusters on average and a maximum of 1000 clusters in each STTpp crate. Every cluster position requires 18 bits (see figure 14) and all clusters correspond to 0.6–2.3 kB of data that have to be read out from the trigger cards.

For each road we also need to save the results of the track fit, which can be characterized for example by

- CFT Level-1 information or track number,
- transverse momentum or radius of curvature,
- azimuthal direction of the track,
- signed impact parameter,
- x position of beam,
- y position of beam,
- χ^2 from fit,
- bit mask that indicates which hits were used in fit.

Each of these should fit into a 16-bit word. For 2–32 tracks that means 32 bytes to 500 bytes of data.

The complete track information is transferred to the CFTpp and read out to Level-3 from there. The serial link transfers data at a rate of 160 Mbits/s=20 MB/s so that this takes 1.6–25 μs. In this scheme the information which roads a cluster lies in would not be saved.

10.2.2 Readout from Trigger Cards only

For diagnostic purposes it is necessary to read out cluster positions along with their associations to roads, but it is probably sufficient to read out the clusters that are in a road. This only requires reading out the processor cards. In this scheme we need to read out 42–500 clusters and 2–32 CFT tracks. This corresponds to $95 + 34 \approx 130$ bytes to $1.1 + 0.5 \approx 1.6$ kB of data

We may also read out the stereo clusters from the trigger cards which amounts to another 0.3–1.1 kB.

If all information is read out from the STTpp crate to Level-3, only "good" tracks have to be transferred from the processor cards to the CFTpp. We define "good" tracks as having hits in at least three layers of the SMT. This reduces the number of tracks to half the number of roads. We then have to transfer on average one track per STTpp crate or 32 bytes, which takes 0.8 μs. For the maximum of 32 roads it takes 26 μs.

10.2.3 Readout from CFTpp only

Alternatively, we can transfer the clusters and tracks in all roads to the CFTpp and have Level-3 read it out from there. This would remove the need for any Level-3 buffer controls and VBDs in the STTpp crate. Transferring 0.13–1.6 kB of data takes 6.5–80 μs.

11 Reconstruction of z-Position of primary Vertices

12 Primary Vertex Determination

We present here the possibility to measure the z position of the primary event vertex using hit information from the SMT at Level-2. The four inner barrel sections of the SMT, which cover the range $-25 \leq z \leq 25$ cm around the center of the detector, are equipped with four layers of double-sided detector elements. Layers 1 and 3 have strips that are parallel to the beam direction on one side and strips that form a 90° angle with the beam direction on the other side. The hit information from the 90° strips only can be used for a fast determination of the z position of the primary event vertex because these strips are projective in the r-z plane and no hit information from the axial strips is necessary. Using these detector elements we can measure the z vertex position in the range $-30 \leq z \leq 30$ cm with a resolution of 525 μm.

Knowledge of the z position of the primary event vertex would allow a more precise measurement of transverse momentum (p_T) at Level-3 from the calorimeter, which measures energy and needs the z vertex position to convert this into a momentum vector. A better transverse momentum resolution translates into reduced trigger rates because the background p_T spectra, which dominate the trigger rate, typically fall rapidly with increasing p_T and because a sharper threshold allows higher threshold values without loss of efficiency. Moreover the knowledge of the vertex multiplicity can be used to veto on events with many interactions, as an intelligent prescale for high rate processes.

In the following we describe the algorithm used, and present a conceptual design of a possible implementation of the algorithm in the trigger hardware. A simulation of the performance and physics benefits from this device is described in reference [15].

12.1 Algorithm

Given the hits in layers 1 and 3 of the SMT we first considered connecting all hits in layer 1 with all hits in layer 3 and histogramming the intercepts of all these lines with the beam axis. Peaks in this histogram would indicate event vertices. However the number of combinations may be prohibitively high (over 4000 for events with two additional interactions) and therefore such an algorithm, which requires many floating point operations, would far exceed the available processing time at Level-2.

We therefore designed an algorithm which is in principle equivalent to the above but utilizes massive parallelism to compute the results within the prescribed time-frame. The loop over the hit combinations is replaced by a loop over candidate vertex locations. The algorithm sequentially assumes discrete values of the vertex z position and for each value it counts the number of hit pairs in layers 1 and 3 that line up. This can be easily achieved by scaling the z positions of all hits in the first layer as follows,

$$z'_{1,i} = \frac{R_3}{R_1} z_{1,i} + \left(1 - \frac{R_3}{R_1}\right) z \tag{5}$$

where $z_{1,i}$ is the i^{th} hit in layer 1, z is the assumed vertex position and R_j is the radial position of layer j. If multiple coulomb scattering is negligible, then $z_{3,i} = z'_{1,i}$ for tracks that point to the assumed z vertex position. Therefore, we can simply bin the hit positions in layer 3 and the scaled hit positions in layer 1 with a resolution that is similar to or coarser than the hit resolution of the silicon devices in z direction (about 30 μm) and count how many bins have entries both from layer 1 and layer 3. Equation 5 can be evaluated for $z = 0$ and histograms for other values of z can be obtained simply by shifting this histogram.

Thus, this algorithm has the advantage that the computation is independent of the hit multiplicity in the event and it only involves one multiplication per hit. Figure 15 illustrates this concept.

A simple peak finding algorithm is used to determine the vertex location. The peak finder scans the track multiplicity for each z-position and identifies the one with largest track multiplicity as the event vertex. To compute the z-vertex a weighted average of the two neighbouring bins is used. The peak finder sets these bins to zero and goes on to find the next highest bin and so on to determine the location of other vertices in the event until the track multiplicity of the highest bin falls below a threshold. This threshold is set to 5 times the average track multiplicity of the event.

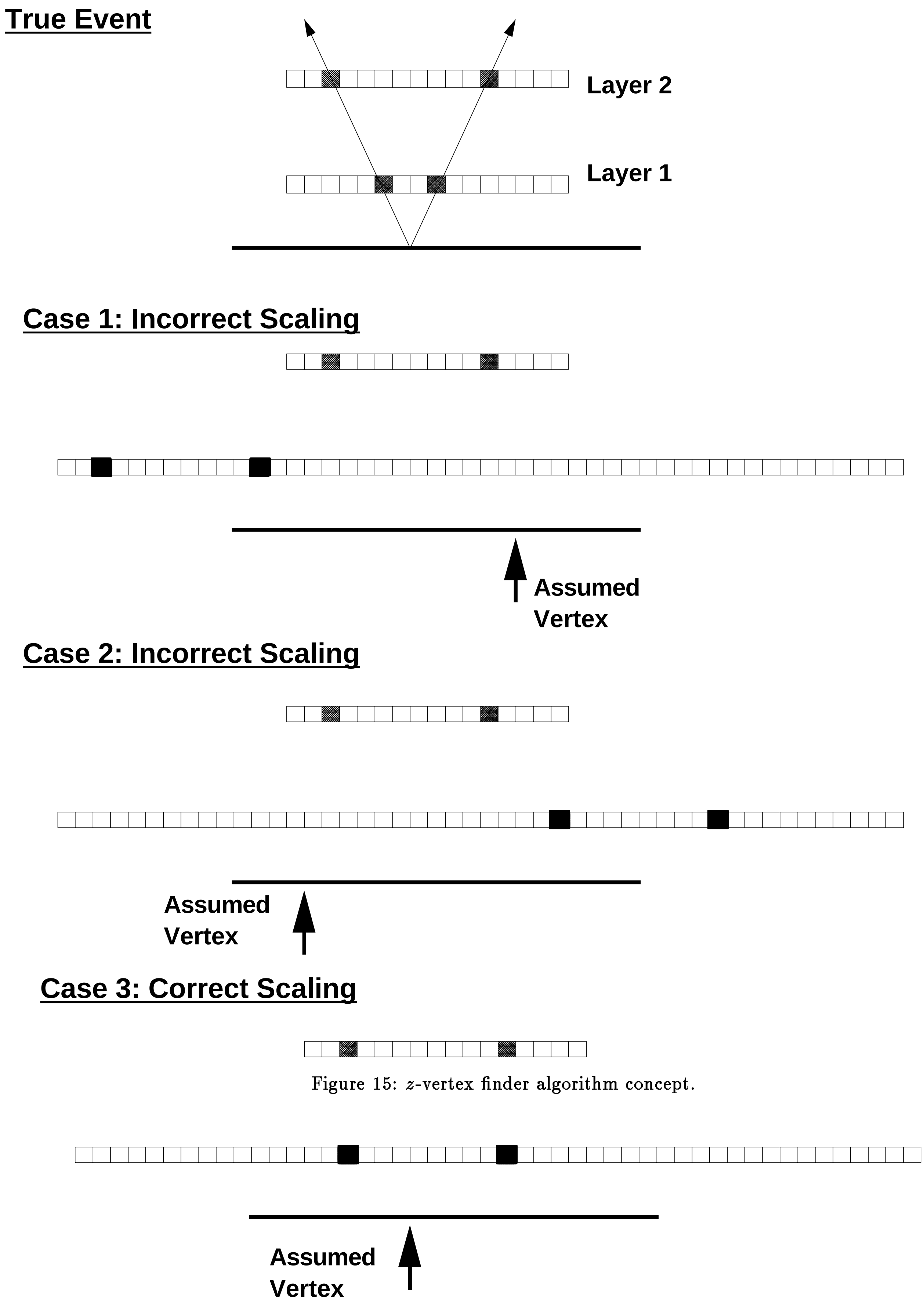

Figure 15: z-vertex finder algorithm concept.

Table 4: Occupancies (in clustered hits)

Number of hits in 30° ϕ sectors			
	1 interaction	3 interactions	6 interactions
hline layer 1	6	13	26
layer 3	4	10	17

12.2 Implementation

We implemented the above algorithm using 1 mm bin sizes for layer 3. We then scaled the hits in layer 1 to the average radius of layer 3 at 7.5 cm. As the even and odd ϕ ladders in layer 3 are at different radii, for simplicity we also scaled the layer 3 hits to 7.5 cm, before binning them into 1 mm bins. A finer granularity (about a factor of 5–10 more) could be used, but is more expensive and time consuming. The granularity of 1 mm gives us acceptable vertex resolution, as will be shown below, and is cost effective. The step size chosen for the vertex scan is also set to 1mm.

We have studied the clustered hit occupancy in these ladders. The hit occupancy (without multiplexing in the 90° ladders and 156 μm strip pitch) in 30° ϕ sectors in layers 1 and 3 for events with one, three, and six interactions is given in table 4.

12.3 Design of the z-vertex finder

For the two-layer system in D0, the tracking problem can be reduced to a simple coincidence logic as described in section 1. The hits in the outer layer are projected to a common radius of 7.5 cm and binned. An adequate bin size, determined from simulation, is 1 mm but the design is flexible enough to accommodate any optimized value. The theoretical lower limit of 35 microns is probably excessively precise but an intermediate value may prove to be more cost-effective.

The region of interest along the beamline is divided into virtual bins, with appropriate bin sizes. Again, a value of 1 mm has been used in this exercise. Now, the hits in the inner layer are scaled as described in equation 5. It can be seen that for the correct value of z, an equality $z'_{1,i} = z_{3,i}$ is formed. By counting such coincidences, the vertex can be located above the background from random coincidences. This transformation consists of a simple scaling of all hits (first term) followed by consecutive shifts by a constant number of bins (for z changing in integral amounts).

The SMT geometry naturally divides itself into 12 overlapping sectors. The z vertex finder takes full advantage of this and performs its computations on 12 cards in parallel. One such sector is shown in figure 16.

The first task is to form centroids using raw data. This function is carried out on the trigger cards. The output of the cluster finder is buffered internally in a DSP, which also scales the hits for layer 1 by the factor appropriate for this sector. The scaling factor will be determined from off-line alignment programs after detector turn-on. The DSP should be able to keep up with the output of the cluster finder. The output of the DSP is a 10-bit number for layer 1 but only 8-bits for layer 2. These hits are decoded and the corresponding bins are latched in large arrays as shown.

So far, all of the functions are carried out in a pipeline and hence the estimated elapsed time is about 10 μs after a Level-1 accept. Now, the hits from the two layers are presented to a parallel coincidence and adder circuit. We estimate that the task requires four FPGAs. The total input pin count is 1024 (256×2 for layer 3 and 256×2 for layer 1). Each FPGA will use 256 pins for the input and 6–8 pins for the output.

The gate count for the coincidence (including neighbor logic) is 2048 and about 4000 for the parallel adder. This will not only fit quite easily into various commercial parts but given the small gate count, it will be able to perform this function in one clock cycle. The outputs of all four FPGAs is added by a simple adder and loaded into a histogram memory. These last two blocks can also be implemented in another FPGA.

Next, the hits in layer 1 are shifted by 3 bins (corresponding to the scale factor) and again fed into the coincidence FPGAs. For a total of 512 z bins, this stage requires 512 clock ticks to fully load the histogram memory. Assuming a 53 MHz clock, this takes about 10 μs. Hence, the total time spent since the Level-1 accept is now 20 μs.

The histograms from all 12 sectors have to be now added. A simple check of bins above some threshold count can be used to define vertices. However, with the remaining approximately 30 μs of the assigned Level-2 preprocessor time, a more sophisticated peak finding algorithm can also be implemented.

13 Beam Alignment and Stability

13.1 Initial Alignment of SMT with Beam

13.1.1 Introduction

One of the purposes of the STT is to allow triggering on tracks originating from displaced secondary vertices, *i.e.* tracks with large impact parameters. Since no determination of the primary vertex can be done at Level-2 time, the impact parameter is determined with respect to a nominal beam position. A misalignment of the detector and beam axes will cause a fake impact parameter, leading to increased background trigger rates and changes in the trigger efficiency. For optimal operation of the trigger, the relative alignment of detector and beam center should be better than 500 μm , and the angle between the detector axis and the beam direction should be smaller than 100 μradians. To achieve this, several actions need to be taken:

- the nominal beam position and direction at DØ expected for run 2 has to be known with high precision.
- the initial installation of the detector has to be close enough to the nominal beam position and direction so that the beam can be steered to conform with the requirements for good operation of detector and trigger.
- to allow compensation for beam drifts during the run, DØ must provide a monitoring system that determines the lateral beam position and tilt and feeds this information to a beam control system provided by the accelerator division.

In Run I, the detector was aligned to less than 2 mm from the beam center in both x and y, and the misalignment in angle was about 1 mrad. In Run II, it is expected [16] that the beam can be moved by at most $\pm$1mm, and the angle can be changed by at most 1 mrad. This is the ideal case, assuming the detector is installed so that one starts from the middle of the range. In addition to this, the beam can also be steered by the correction magnets D11 and C49. The range of beam steering using these magnets is about 100μm in position and 100μrad in angle. If we assume that it will be the same in Run II, then the initial placement of the detector has to be much more precise than was achieved in Run I. If this is not possible, we have to get the accelerator to extend the range of possible beam motion.

13.1.2 Action Plan for DØ

In order to make sure that we have a fair chance to install the detector with the best possible initial alignment with respect to beam position and direction, DØ should survey the beam during the collider commissioning period, similar to the procedure envisaged by CDF. The accelerator seems in fact to expect such a request from DØ. It may also be useful for the accelerator to know where the beam is at DØ. The collider commissioning period may be long enough to do these measurements after CDF has done them, *i.e.* it is probably possible to share the equipment with CDF. The equipment needed consists of BPM's (beam position monitors) and flying wire assemblies (see fig. 17). CDF's goal is to measure the beam position to better than 0.1 mm and the beam angle to within 100μrad, relative to survey targets on the wall of the collision hall and on the low beta quadrupoles. This will allow to correct for possible sinking of the floor of the collision hall after the installation of the detector. The setup and goals of these measurements are described in a CDF note [17].

We have to investigate whether the present survey system at DØ is sufficient to provide alignment information which meets the precision requirements of the SMT.

We have to start an aggressive R&D program to develop techniques to install the detector to better than 1 mm precision at the nominal beam crossing point, and with axial alignment better than 0.5 mrad.

Given the size and mass of our central detector, it is probably not realistic to assume that a system for remote-controlled motion of the detector could be designed to move the detector to its nominal position

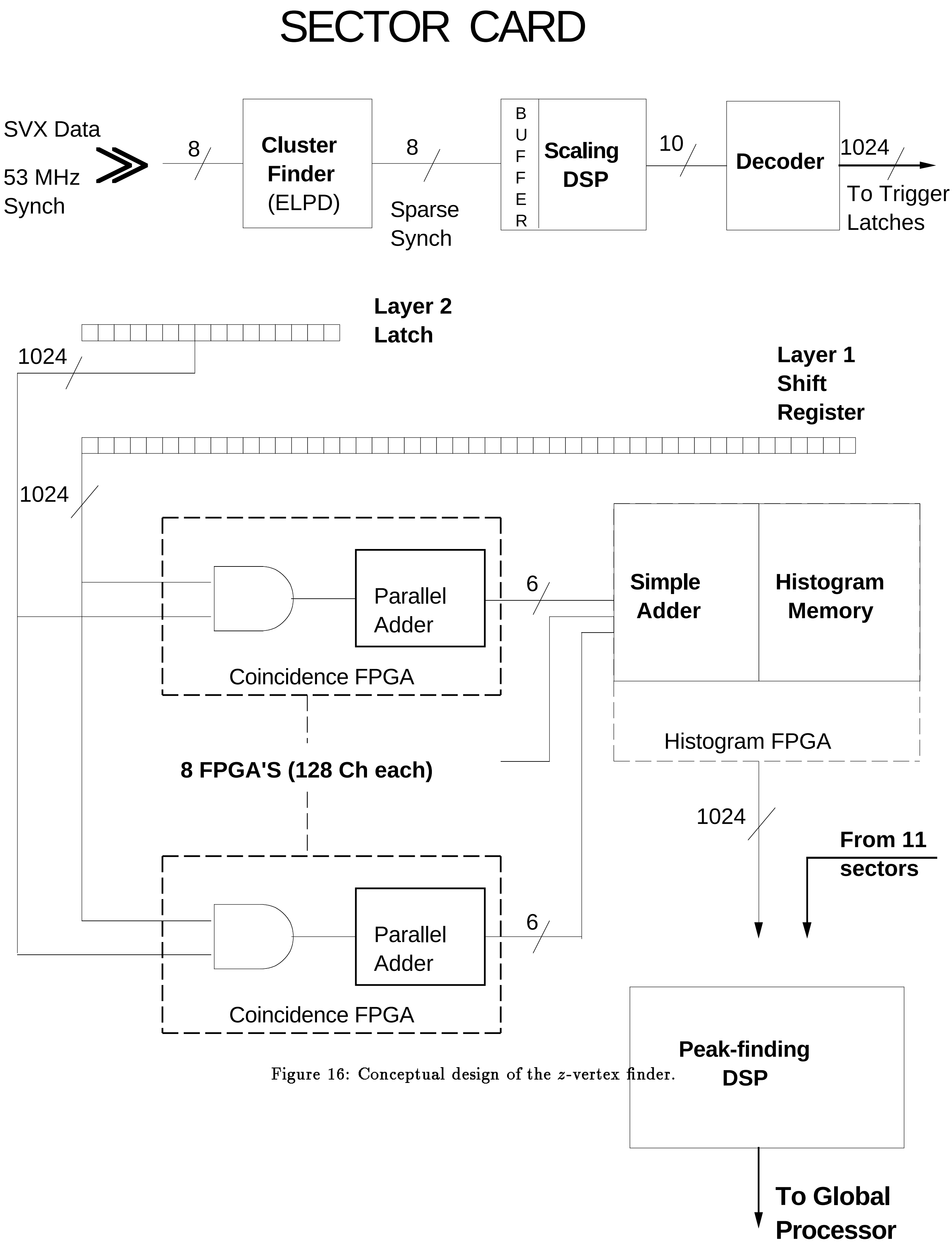

Figure 16: Conceptual design of the z-vertex finder.

if the initial installation is not sufficiently precise. Also, motion may cause stress leading to unknown and uncontrollable distortions whose harmful effects may more than compensate for the benefits of better alignment.

We have to increase our presence in the mind of the beam division personnel, by being more aggressive in our requests for support.

13.2 Beam Stability

13.2.1 Experience from Run I

Variations in beam position
During Run I, CDF monitored the beam position using their silicon vertex detector. The silicon vertex detector data were used to reconstruct track parameters, including the primary vertex position. In Fig. 18 (for run IA) and Fig. 19 (for Run IB), the centroid position and angle between the detector axis and beam axis for both horizontal and vertical planes are shown. The CDF observations can be summarized as follows: The average displacement was about 0.5 mm, and the average angle between beam and detector axis was $\approx 500\mu$rad. There was a long-term drift in both position and angle, with discrete steps before and after shutdown periods. Typical store to store variations in beam centroid position were $\approx 100\mu$m horizontal and $\approx 150\mu$m vertical. The 2000 μm step in the Run IB horizontal position occurred when the B0 quadrupole roll was fixed.

Angle variations from store to store were typically less than 100μrad, but of the order of $\approx 200\mu$rad throughout the runs. On the other hand, changes of position during the store were much less important, typically $\approx 40\mu$m horizontal and vertical, while angle variations during a store were typically too small to be measurable. The reasons for the store-to-store variations are not quite understood (quad motion, beam steering,..??).

Feedback system for beam position control
In order to reduce the variations in beam position and stability, CDF in collaboration with the accelerator division, implemented a feedback beam control system, using the measured beam position to set dipole corrector magnets near B0 to stabilize beam position and angle [18]. The *maximum possible range* of variation is determined by the range of the magnet power supply ($\pm$50A), corresponding to a maximum angle correction of $\pm 100\mu$rad However, the *practical range* of variation is limited by good field region requirements and aperture limitations, and therefore the actual range is smaller than this, the precise value being unknown. Tests performed by CDF and the accelerator division in June 1995 showed that, with 5 minute updates, the beam position could be kept stable to within $\approx 5\mu$m, and the beam angle to within $\approx 50\mu$rad. In a further test in February 1996 the beam was moved 50μm vertically, 50μm horizontally in the B0 interaction region. Closed orbit measurements verified that the bumps were local to the B0 interaction region *without change elsewhere around ring*, an indication that they can control beam position at DØ and B0 independently.

13.2.2 Expected Variations during Run II

No clear information available so far.

13.3 Alignment Stability and Maintenance

13.3.1 Measurement of Beam Trajectory

The information that needs to be provided to the beam position control system consists of four numbers: (X_0, Y_0), the x, y position of the interaction point centroid (x = horizontal, y = vertical), and the slopes A_x, A_y of the line of the interaction point centroids versus z, in the xz and yz planes. These numbers can be determined by using the tracker information to reconstruct tracks coming from the primary vertex. The reconstruction yields for every track the following information: curvature, polar angle θ, impact parameter b (*i.e.* distance of closest appproach to the detector axis, *i.e.* the line $(x = 0, y = 0)$), φ_b (azimuth of point of closest approach), and z_b (z of the point of closest approach). If the beam is perfectly aligned with the axis of the detector, *i.e.* centered on the line $x = 0,\ y = 0$, the impact parameter for tracks coming from the

tilt	< imp.par. >	fraction (%) above		
(μrad)	(μm)	100	200	400
0	32	1	0	0
100	35	2	0	0
200	43	5	0.1	0
400	66	20	2	0
600	92	38	8	0
800	118	51	17	0
1000	146	58	27	3
2000	218	77	57	26

Table 5: Fake impact parameter due to beam size and misalignment. For various beam tilt angles, the average fake impact parameter is given, as well as the fraction of tracks with impact parameter bigger than 100, 200, 400 μrad, assuming the impact parameter is calculated with respect to the line ($x = 0, y = 0$).

primary interaction point is only due to the finite transverse beam size, and there is no correlation between b, φ_b and z_b. If the beam is misaligned in position and/or angle, then there is an additional contribution b_{mis} to the impact paramater which is a function of φ_b and z_b:

$$b^2 = b_s^2 + (b_{mis}(\varphi_b, z_b))^2$$

$$b_{mis}(\varphi_b, z_b) = X_0 \cdot sin\varphi_b + A_x \cdot z_b \cdot sin\varphi_b - Y_0 \cdot cos\varphi_b - A_y \cdot z_b \cdot cos\varphi_b,$$

where X_0, Y_0 = center of beam at $z = 0$, A_x, A_y = x and y slope of beam, b_s^2 = contribution to impact parameter due to finite beam size ("transverse beam smearing").
For a set of N tracks with measured impact parameters b_i $(i = 1, N)$, which come from the primary vertex, a minimization of

$$\chi^2 = \sum_{i=1}^{N} (\frac{b_i - b_{mis}(\varphi_{bi}, z_{bi}|X_0, Y_0, A_x, A_y)}{\sigma_i})^2$$

with respect to the parameters (X_0, Y_0, A_x, A_y) will yield estimates of the beam parameters needed for the beam control feedback system. This method of determining the beam position and direction has been used by CDF in Run I [19, 20].

13.3.2 Requirements on Relative Alignment of Detector and Beam

In the vicinity of the interaction point the beam has a Gaussian width which varies with the distance from the nominal crossing point ($z = 0$):

$$\sigma(z) = \sigma(0) \cdot \sqrt{1 + (\frac{z}{\beta^*})^2},$$

where $\sigma(0)$ is the beam size at the nominal interaction point, and z is the distance from the nominal IP. With the Run II values of ϵ_{proton} 3D = 30πmm mr and $\beta^* = 35$cm, σ_0 is expected to be $= 40.5\mu$m. Figures 21 to 27 show the effect of beam tilt on the impact parameter distribution, assuming the beamsize as given above, and that the z distribution of the primary vertices is a Gaussian with $\sigma_z = 25$cm. Table 5 summarizes the effect of beam tilt. For beam tilt angles larger than $\approx 200\mu$rad, the fraction of fake impact parameters becomes unacceptably large, and would give rise to unwanted triggers. Using the barrel midpoint as a reference to calculate the impact parameter would alleviate this effect.

A study of the effect of the finite beam size and a beam position shift on $t\bar{t}$ events [21] showed:

- the finite beamsize has a small effect: for a fixed background rate, it leads to a small (at most) 5% reduction of the signal efficiency (see fig. 28)

- a systematic shift in the primary vertex position will limit the expected reduction of the background rates but will not greatly affect the signal efficiency (see fig. 29). In particular, in order to keep background rates under control, one should try to keep the beam shift below $\approx 30\mu$m.

Note that a beam tilt of 200μrad will cause an effective shift of the beam by $\approx 24\mu$m over the length of a SMT barrel. We conclude that if we use the barrel information to calculate the impact parameter with respect to the beam centroid at the z corresponding to the barrel midpoint, beam tilts up to $\approx 200\mu$rad can be tolerated.

13.4 Feedback Mechanism to Accelerator

In order to compensate for beam position and angle drift during a store, CDF in collaboration with the Beam Division, has developed a beam monitoring and control system. This system measures beam position and slopes in real time and makes this information available to the accelerator controls system (ACNET). The system is described in CDF note 3859 [18].

We propose to implement a similar system for DØ. The determination of the beam parameters (X_0, Y_0, A_x, A_y) could be performed by a dedicated processor at Level-3, using the procedure outlined in sect. 13.3.1.

Since the algorithm is based on tracks rather than events, the number of events is expected to be rather small — a few hundred should be ample. The events subjected to this should be preferably events which have not been obtained by selecting high impact parameters, although this is not really a crucial requirement. For every track, the track parameters b, ϕ_b, z_b will be stored in an array. After having accumulated sufficiently many tracks (about 500), a fit is performed to determine the estimates for the beam parameters. From the beam parameters, and z_b, the ccordinates of the primary vertex can be calculated for every track. This allows the calculation of a corrected impact parameter b_{corr} which is the impact parameter with respect to the primary vertex, rather than with respect to $(0, 0, z_b)$. The tracks with corrected impact parameter above some threshold are then discarded and the fit is repeated. The total number of tracks to be used in the fit, a possible p_t threshold value, the cut-off for the corrected impact parameter and the number of iterations will have to be determined by further studies. From the CDF Run I experience, it appears that applying the beam correction every few minutes is enough to keep the beam position and angle sufficiently stable, *i.e.* there is no problem with computing time.

The results of the fit, *i.e.* the beam parameters, are then fed to the beam control system. The ACNET devices for this exist already: C:D0XIP, C:D0YIP $= x, y$ positions of beam, *i.e.* X_0, Y_0; C:D0DX, C:D0DY $=$ slopes in xzandyz plane, *i.e.* A_x, A_y. There are other ACNET devices relating this, *e.g.* beam size in x, y, run number, and time stamps of the various parameters. The parameters submitted to ACNET from the experiment are then used by existing ACNET applications to change settings in correction magnets, in order to maintain stable beam position and angle.

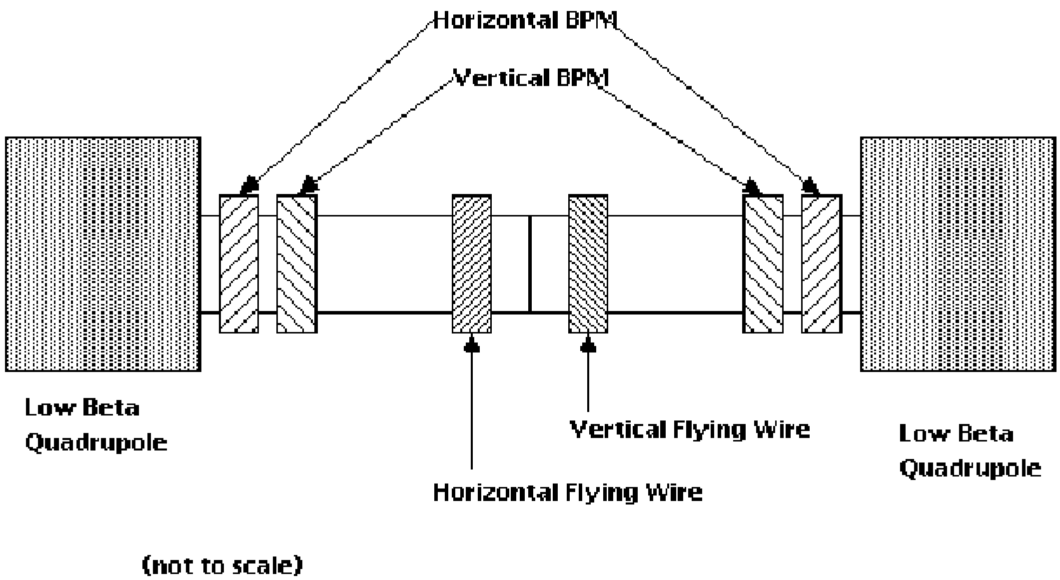

Figure 17: CDF's proposed setup to determine the initial beam position for Run II. (Figure from ref. [17])

14 Simulation of Performance

See physics study notes [24].

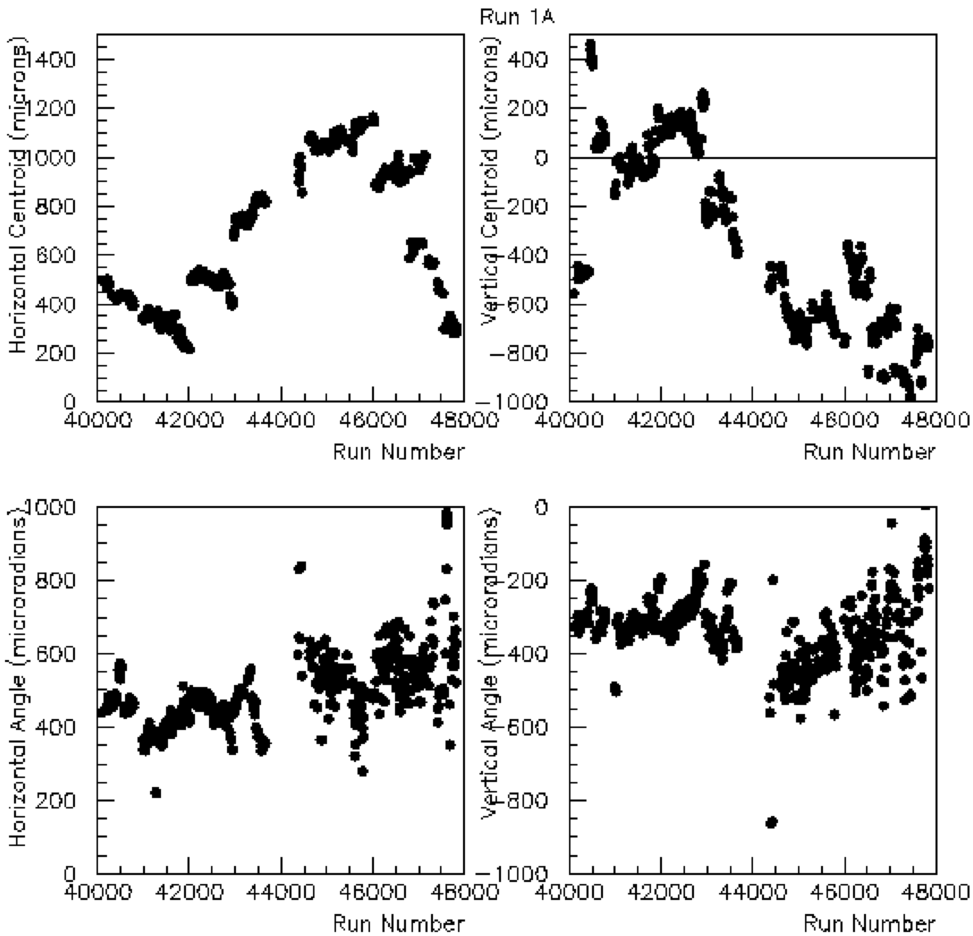

Figure 18: Horizontal and vertical beam positions and angles for Run IA, as measured by CDF. Every entry corresponds to one data-taking run.

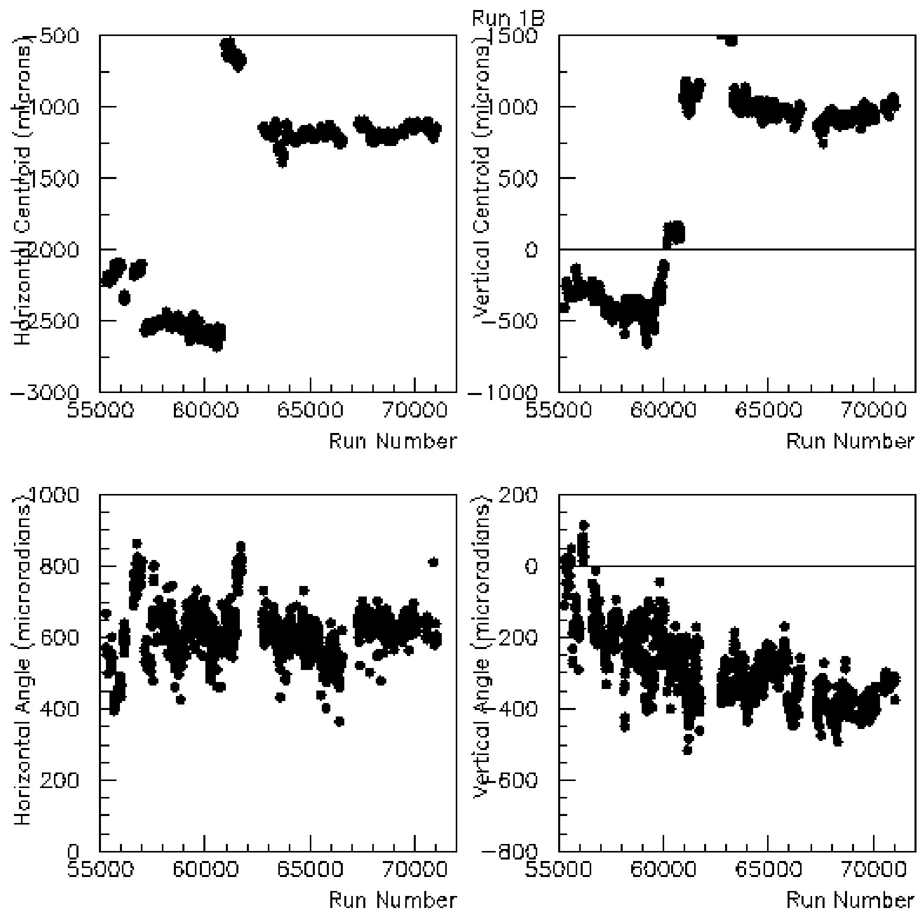

Figure 19: Horizontal and vertical beam positions and angles for Run IB, as measured by CDF. Every entry corresponds to one data-taking run.

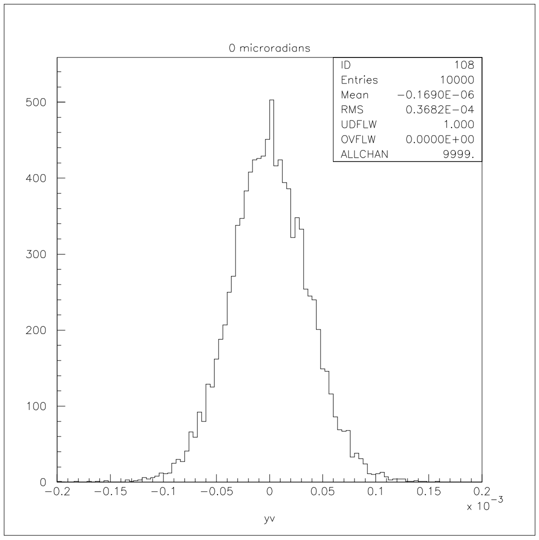

Figure 20: y(vertex) distribution for zero tilt angle

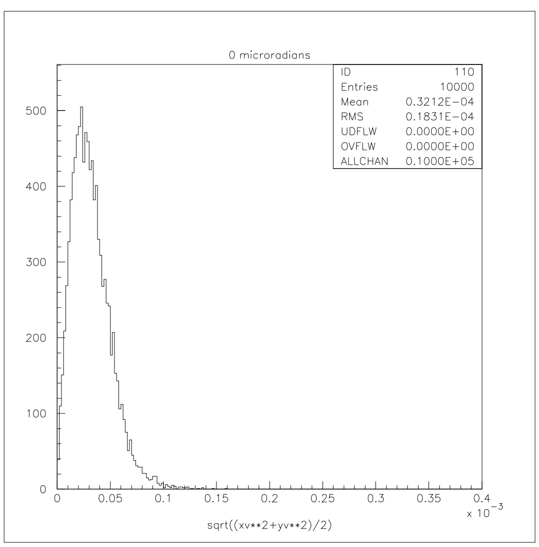

Figure 21: Fake impact parameter distribution with respect to (0,0) for tracks coming from primary vertex, for zero tilt angle.

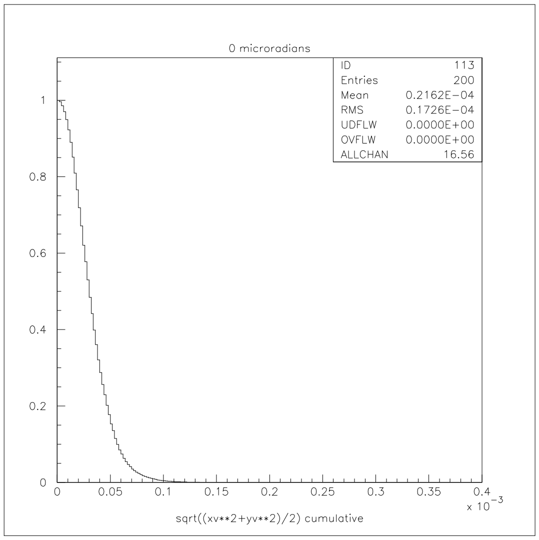

Figure 22: Normalized cumulative distribution of fake impact parameter with respect to (0,0) for tracks coming from primary vertex, for zero tilt angle.

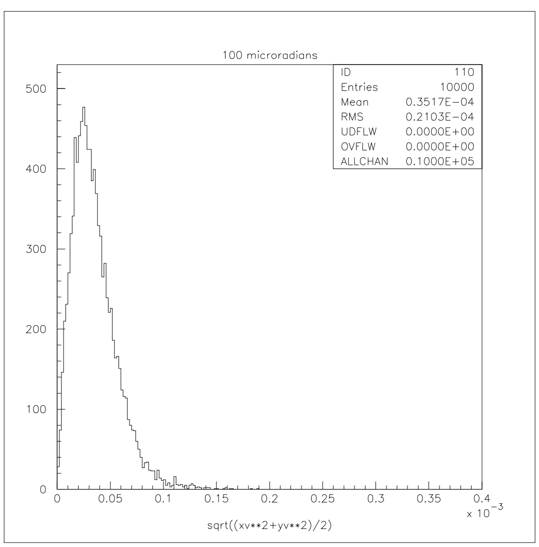

Figure 23: Fake impact parameter distribution with respect to (0,0) for tracks coming from primary vertex, for tilt angle = 100μrad.

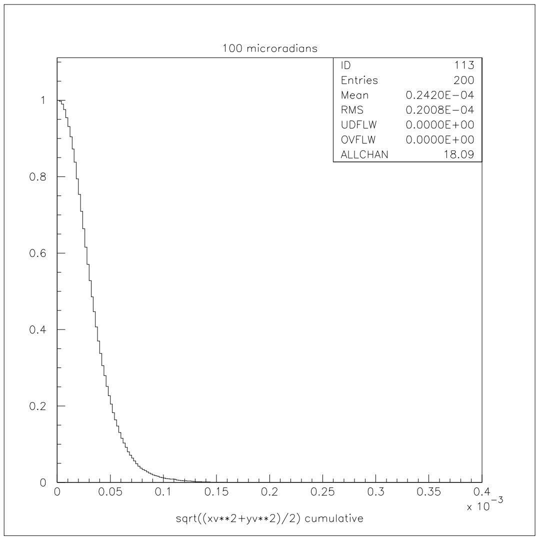

Figure 24: Normalized cumulative distribution of fake impact parameter with respect to (0,0) for tracks coming from primary vertex, for tilt angle = 100 μrad.

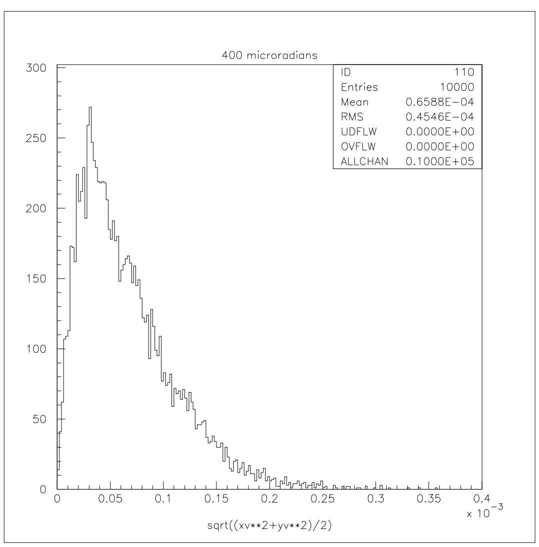

Figure 25: Fake impact parameter distribution with respect to (0,0) for tracks coming from primary vertex, for tilt angle = 400μrad.

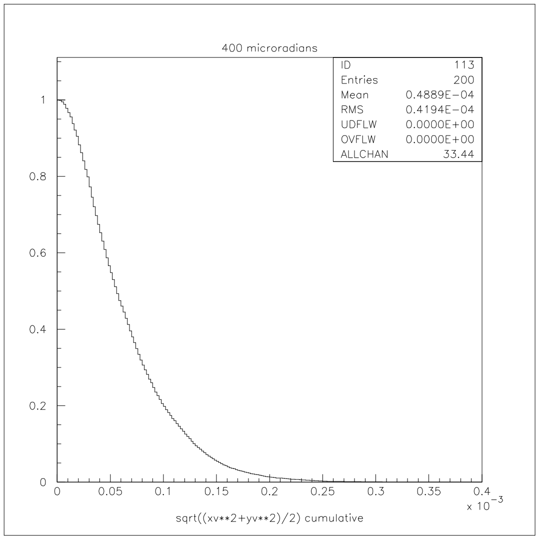

Figure 26: Normalized cumulative distribution of fake impact parameter with respect to (0,0) for tracks coming from primary vertex, for tilt angle = 400 μrad.

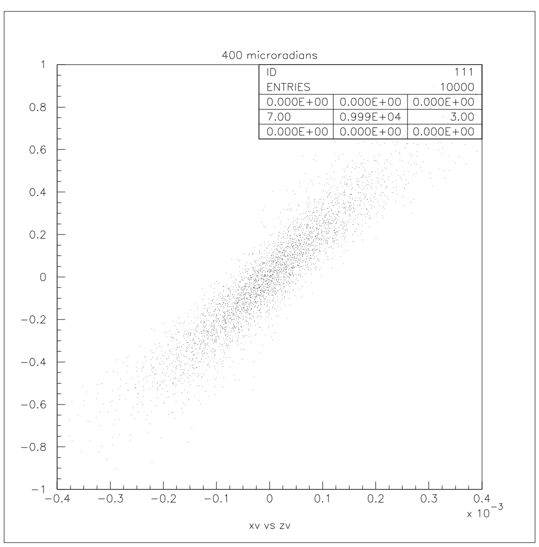

Figure 27: Distribution of x(vertex) vs z(vertex), for 400μrad tilt angle in xz plane.

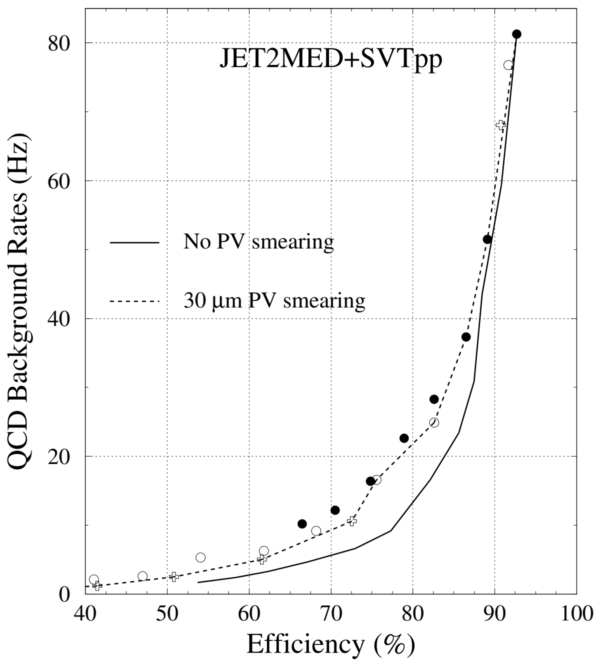

Figure 28: Signal efficiency for $t\bar{t} \to$ all jets versus QCD background rate, with (dashed line) and without (solid line) smearing of the primary vertex.

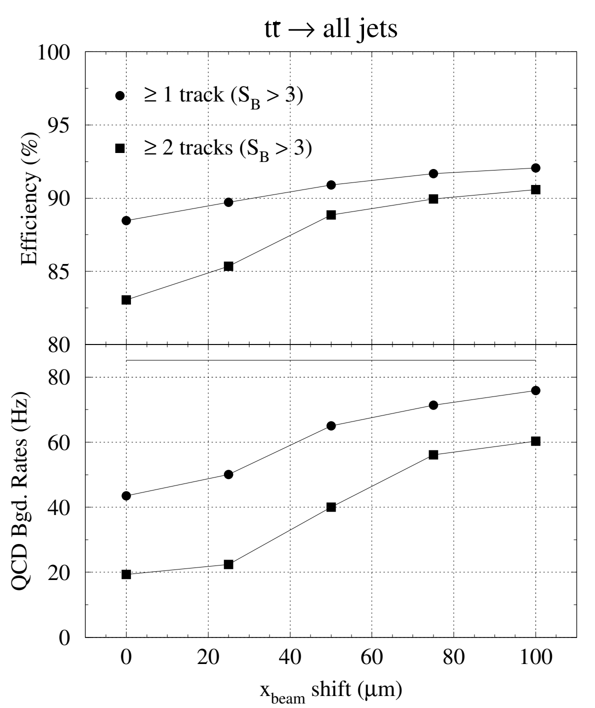

Figure 29: Signal efficiency for $t\bar{t} \to$ all jets versus QCD background rate, vs beam shift in μm, for two different trigger conditions: impact parameter significance $S_b > 3$ for at least one track (solid circles), and for at least two tracks (solid squares).

15 Cost Estimate and Schedule

Table 6 shows the cost estimate for development and construction of the full system. The project requires 5 person-years (py) of engineering support and \$ 740k in equipment funds.

Table 6: Cost estimate for STTpp project.

Item	engineering	unit cost	number	total cost
trigger card	1.5 py	\$ 5200	60	\$ 312k
FPGAs for trigger card	0.5 py	\$ 273	500	\$ 137k
processor card	1.5 py	\$ 10000	12	\$ 120k
receiver/controller card	0.5 py	\$ 5200	6	\$ 31k
transceiver card[1]	—	\$ 5000	8	\$ 40k
alpha processor[1]	—	\$ 15000	2	\$ 30k
VME crates/power supplies	—	\$ 10000	7	\$ 70k
system integration	1.0 py	—	—	—
total project cost	5.0 py			\$ 740k

[1] under development for the DØ trigger system

Table 7 shows a schedule for the design, testing, implementation and commissioning of the STTpp. The device will not be available for the initial commissioning phase of Run II, which starts in April 2000. At the beginning of the run the luminosity is expected to be below design and the rejection afforded by the STTpp is not required. The full impact of the STTpp will come to bear later in the run, when the accelerator reaches its design luminosity.

Table 7: Schedule for development of STTpp.

date		milestone
April	1998	collaboration review of STTpp design
September	1998	award of grant by NSF
June	1999	prototype subprocessor crate
October	1999	finalize design, begin production
March	2000	system test
April	2000	begin Run II
August	2000	installation at DØ
October	2000	STTpp operational

16 Conclusion

We have presented an initial design of a Silicon Track Trigger preprocessor for the DØ Level-2 trigger system in Run II. This processor will associate hits in the SMT with CFT tracks, found at Level-1 and fit them jointly for precise impact parameter and curvature determinations. Knowledge of the impact parameter at Level-2 will allow more efficient data collection for physical processes with b quarks in their final state. Such processes are of fundamental importance to many of the questions that we want to address in Run II. The proposed processor also offers the possibility to compute the position along the beam trajectory of the primary $p\overline{p}$ interactions in an event. This information is important to improve momentum resolution for electrons and jets in Level-3 and can serve as seed information for a fast tracking algorithm in Level-3. The design presented shows that implementation of this functionality is possible with commercially available components. Its details will have to be worked out in the coming months.

Acknowledgements

We would like to acknowledge the help of J. Blazey, F. Borcherding and M. Martin in understanding the CFT Level-1 trigger and S. Mishra in obtaining accelerator information.

References

[1] S.L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, *Proceedings of the 8^{th} Nobel Symposium*, ed. N. Svartholm (Almqvist and Wiksells, Stockholm 1968), p. 367.

[2] "A Measurement of the W Boson Mass", B. Abbott *et al.*, FERMILAB PUB-97/422-E, submitted to Phys. Rev. Letters and FERMILAB PUB-97/423-E, submitted to Phys. Rev. D.

[3] D. Ward, EPS, Jerusalem, 1997.

[4] S. Abachi *et al.*, Phys. Rev. Lett. **79**, 1197 (1997);
S. Abachi *et al.*, Phys. Rev. Lett. **80**, 2063 (1998).

[5] "Measurement of the Top Quark Mass", F. Abe *et al.*, FERMILAB-Pub-97/284-E, submitted to Phys. Rev. Lett..

[6] D. Amidei and R. Brock, eds., FERMILAB-PUB-96/082.

[7] E. Eichten, K. Lane, J. Womersley, Phys. Lett. **B405**, 305 (1997).

[8] M. Johnson *et al.*, DØ note 3168; "Preliminary Hardware Design for a Silicon Vertex Trigger", M. Johnson, February 1998.

[9] "Road Definition for Silicon Track Trigger", DØ note 3363, U. Heintz, 1998.

[10] reference for VRB.

[11] reference for VTM receiver module.

[12] "Dataflow Requirements for the Silicon Track Trigger", DØ note ????, M. Narain, 1998.
The expected averages are based on an average of 4 interactions/crossing, corresponding to $\mathcal{L} = 2 \times 10^{32}$ cm^{-2}s^{-1} and a crossing interval of 396 ns. The expected maximum is based on the endpoint of the spectra for 8 interactions.

[13] `http://d0sgi0.fnal.gov/~foco/vertex/tbtalk.ps`

[14] `http://www.pa.msu.edu/hep/d0/l2/mbt.htm`

[15] DØ note 3338, M. Narain, S. Mani, 1998.

[16] John Marriner, private communication.

[17] "Interaction region measurements during run II startup", Paul Derwent, Brenna Flauger, CDF note 4059, March 1997.

[18] "Beam Motion and Feedback", Paul Derwent, CDF note 3859, August 1997.

[19] "Fitting the beam position with the SVX", Hans Wenzel, CDF note 1924, 1993.

[20] "The primary interaction vertex", Hans Wenzel, CDF note 4066, 1997.

[21] "Effect of primary vertex uncertainty on SVTpp efficiency", Frédéric Stichelbaut, December 1998, (`\http://ww-d0.fnal.gov/\~wahl/minutes/`).

[22] "A preliminary consideration of hit finders for the DØ impact parameter trigger", Henryk Piekarz, DØ note 3209, March 1997.

[23] "The run II handbook", Fermilab Beam Division, `http://www-bd.fnal.gov/lug/runII_index.html`.

[24] "Benefits of a Silicon Vertex Trigger for ttbar Events", F. Stichelbaut, DØ note 3312, September 1997;
"Studies of SVT Trigger using SUGRA Four b Quark Signature", D. Norman, DØ note 3318, September 1997;
"Towards a Run II Trigger for $p\bar{p}$ to W Higgs", J. Hobbs, DØ note 3321, October 1997;
"Triggering on $Z \rightarrow b\bar{b}$ Events in Run II", M. Narain, DØ note 3339, October 1997.