DØnote # ????

A Silicon Track Trigger for the DØ Experiment in Run II

— Proposal to Fermilab —

The DØ Collaboration

September 17, 1998

Contents

1 Introduction

1.1 Purpose of this Document

This document describes the design of a Silicon Track Trigger preprocessor (L2STT) for the DØ detector in Run II [1]. The L2STT interfaces with the approved trigger system of the DØ detector and requires no modification of the approved system. The L2STT allows the inclusion of the last major detector system into the DØ trigger scheme, approved in the baseline upgrade project (joining the central fiber tracker, the preshower detectors, the calorimeters, the muon systems and the luminosity detectors). The L2STTf 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 and τ leptons, at the trigger level and enhances the capabilities of the DØ detector to trigger on charged particles in general. The L2STT improves the capabilities of the DØ detector to exploit the large range of high-p_T physics that will be accessible in Run II.

This introductory section first motivates the addition of such a device based on the goals of the Run II physics program. It then gives a brief description of the components of the DØ detector that the L2STT interfaces with, the trigger system and the SMT. Section 2 describes the conceptual design of the processor hardware and section 3 describes simulation studies of the performance of the device. In section 4 we return to the Run II physics program and use the simulation of the proposed hardware to demonstrate the improvements afforded by the L2STT in triggering on various processes of interest. Section 5 gives a preliminary schedule, cost estimate and funding outline. The document concludes with a summary in section 6.

1.2 Physics Motivation

1.2.1 The Physics Program for Run II

During Run II, the Fermilab Tevatron will run at a center-of-mass energy of 2 TeV and deliver an integrated luminosity of 2–4 fb^{-1}, 20–40 times more than during Run I. The TeV33 program envisions further luminosity upgrades to the accelerator to accumulate some 30 fb^{-1} before the LHC starts operation at CERN. The large data set expected from these runs will allow sensitive studies of the standard model of the electroweak interactions [2]. Among these are intermediate vector boson physics, a detailed study of the top quark, and the searches for the Higgs boson and other new particles, and for supersymmetry or other extensions of the standard model. These studies address some of the most fundamental problems in particle physics research today: the question of electroweak symmetry breaking and the generation of masses. Other fundamental issues of great interest that can be addressed in Run II are the properties of b-hadrons and the search for CP violation in the B meson system.

1.2.2 Studies of the Top Quark

The dominant production process for top quarks in $p\overline{p}$-collisions is $t\overline{t}$-pair production via the strong interaction with a cross section of about 5 pb. Depending on the decay of the W bosons from the $t \to Wb$ decays we distinguish the dilepton channel ($t\overline{t} \to \ell^+\nu b\ell^-\overline{\nu}\overline{b}$), the lepton+jets channel ($t\overline{t} \to \ell\nu bq\overline{q}\overline{b}$), and the all-jets channel ($t\overline{t} \to q\overline{q}bq\overline{q}\overline{b}$). In a data set of 4 fb^{-1}, we expect to identify several thousand lepton+jets events. This sample of $t\overline{t}$ 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 [3, 4, 5] and the top quark mass [6, 7]. 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 large value. With a mass of about 175 GeV, the top quark is the heaviest known fundamental particle and it may well hold the key to understanding electroweak symmetry breaking and the generation of the fermion masses.

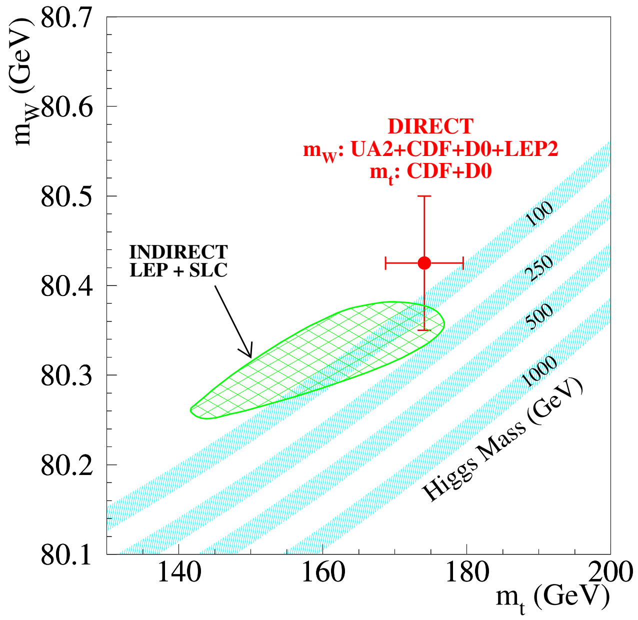

Figure 1: Predicted W boson mass versus the top quark mass.

The most precise top quark mass measurements from Run I use 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 can be identified, this reduces to a two-fold ambiguity, significantly reducing the associated systematic uncertainties. With double b-tagging, the jets from the hadronic W boson decay are identified uniquely and the $W \to q\bar{q}$ signal can be reconstructed in $t\bar{t}$ events without combinatoric confusion.

The $W \to q\bar{q}$ signal would give a precise calibration of the energy scale for light-quark jets. To further reduce the jet energy scale uncertainty, b-quark jets also have to be calibrated more precisely. This can be achieved using Z bosons that decay to $b\bar{b}$. The signature for this process is two jets and therefore swamped by gluon-jet background. At a luminosity of 2×10^{32} cm^{-2}s^{-1}, the event rate for a trigger requiring ≥ 2 jets with $p_T > 15$ GeV is almost 300 Hz. Raising the p_T threshold would begin to cut into the acceptance for the Z resonance. Therefore b-jet tagging at the trigger level is needed to reject the gluon-jet background. The CDF collaboration has demonstrated that it is possible to reconstruct a $Z \rightarrow b\bar{b}$ signal provided both b jets can be tagged [8]. With these additional checks on energy scale and combinatoric uncertainties we should be able to measure the top quark mass to a precision of 3 GeV or better in Run II.

It is important to measure the branching ratios for the top quark decay channels in order to verify that there are no unobserved decay channels. This can be done by measuring cross section times branching ratio for the dilepton, lepton+jets, and all-jets channels and taking their ratio to cancel out the production cross section. Recently D0 has measured the $t\bar{t}$ cross-section in the all jets channel [14] with a precision rivalling the combination of all other channels. The tagging of b-quarks is the key to this analysis, as it is absolutely essential in reducing the backgrounds.

Production of single top quarks via the electroweak interaction is also an interesting process to study. The cross section is smaller than for $t\bar{t}$-pair production, about 3 pb, and the signature less striking, so that this process could not be observed in the Run I data. For $m_t = 170$ GeV, Run II with an integrated luminosity of 4 fb^{-1} will produce approximately 12000 single top quarks [9, 10]. Depending on the decay of the W boson from the $t \rightarrow Wb$ decay, these events have $\ell\nu b\bar{b}$ or $q\bar{q}b\bar{b}$ final states, where one of the b quarks can be very soft. Single top quark production proceeds via a Wtb-vertex and its cross section is thus proportional to $\Gamma(t \rightarrow Wb)$ which, in turn, is proportional to $|V_{tb}|^2$. By measuring the production cross section for single top quarks in Run II, V_{tb} can be measured with a fractional precision of 10% [10]. The tagging of b-quarks is the key to clean single top samples, especially in the hadronic decay modes.

1.2.3 Search for the Higgs Boson

The most promising process in which to observe a standard model Higgs boson at the Tevatron is associated production with vector bosons. Although the cross section for $p\bar{p} \rightarrow WH$ is very small (300 fb for $M_H = 100$ GeV), the leptonic decays of the vector boson provide an easy tag for the events. 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 [2]. This mass window is extremely interesting from experimental and theoretical points of view. The current lower bound on the Higgs mass is 89 GeV[11] and fits to the present electroweak data, assuming the completeness of the standard model, give $m_H = 115^{+116}_{-66}$ GeV [12]. 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 this window be covered reliably. LEP2 will not be able to access Higgs masses above approximately 105 GeV and LHC covers this mass region mainly via the loop induced $H \rightarrow \gamma\gamma$ decay. The Tevatron provides the best opportunity to see the primary decay mode of the Higgs boson in this mass range.

To see a Higgs signal during Run II, the inclusion of additional channels, like Higgs bosons produced along with hadronically decaying W or Z bosons, is required to increase the sensitivity beyond the study in reference [2]. For WH production, 67% of the decays are to $q\bar{q}b\bar{b}$. For ZH production 70% of the events decay to $q\bar{q}b\bar{b}$. These events contain four jets, a signature that is swamped by QCD production of gluon and light-quark jets. To keep the jet p_T threshold low we need b tagging in the trigger to reduce the background trigger rate.

To interpret the results of a Higgs search, the observation of a $Z \rightarrow b\bar{b}$ signal is almost mandatory. It would prove that we can see a narrow resonance that decays to $b\bar{b}$ and it would also provide essential measurements of b-tagging efficiency and $b\bar{b}$-mass resolution.

1.2.4 Beyond the Standard Model

Technicolor breaks the electroweak symmetry dynamically by introducing a new strong interaction. It predicts the existence of particles in the TeV mass range, some of which couple to mass, like the Higgs boson, and therefore decay to final states containing b quarks. In $p\overline{p}$ collisions the techni-ρ (ρ_T) and the techni-ω (ω_T) may be produced. If allowed, the ρ_T decays predominantly to technipions which decay to heavy fermions, *e.g.* $\rho_T^+ \to \pi_T^+\pi_T^0 \to b\overline{c}b\overline{b}$. If the mass of the ρ_T is less than twice the mass of the π_T, the decay $\rho_T^+ \to W^+\pi_T^0$ dominates. This process has a signature identical to that of Higgs production in association with W bosons, but a significantly higher cross section (several pb) because of the ρ_T-resonance enhancement, so that it should be visible in Run II [13].

Topcolor explains the high mass of the top quark through new strong dynamics. It predicts Z' bosons and top gluons that preferentially couple to the third fermion generation. Their decays to $b\overline{b}$ would give rise to resonance structures in the $b\overline{b}$-mass spectrum.

Supersymmetry is a symmetry between bosons and fermions, from which the Higgs mechanism can be derived. It requires the existence of superpartners to all known particles. The superpartner of the top quark (the top squark $\widetilde{t}$) is pair produced in $p\overline{p}$ collisions and may decay via $\widetilde{t} \to b\widetilde{\chi}_1^+$. For hadronic decay modes of the chargino $\widetilde{\chi}_1^+$, b tagging in the trigger is needed. If supersymmetry is discovered, the ability to identify b quarks will be of great use in elucidating the sparticle states involved.

1.2.5 b-Physics

Another important field that will be studied with the data acquired during Run II is the physics of b hadrons. During the coming years, b physics will be intensively studied at the e^+e^- B factories at SLAC, KEK and Cornell. The copious production of b hadrons of several species at the Tevatron offers the opportunity to provide competitive measurements with respect to e^+e^- machines. Some of the topics to be studied are:

- the search for CP violation in the B sector with the measurement of the angle β of the unitarity triangle using the golden mode $B_d^\circ \to J/\psi K_s$;
- the search for rare decay modes, like $b \to \mu^+\mu^- X_s$, $B_s/B_d \to \mu^+\mu^-$, or the radiative B decay $B \to \rho\gamma$;
- the measurements of V_{ub} and V_{cb} using the semileptonic decay channels of B mesons.
- the search for B_s mixing above the current limit from the LEP experiments of $x_s > 15$.

The search for CP violation in B mesons decay is a topic of major interest in high energy physics. CP violation is one of the least understood phenomena in the standard model and its understanding will have far-reaching ramifications in cosmology and high-energy physics. Through tagging b-quarks at the trigger level, the L2STT will give improved samples of $b\overline{b}$ final states faciliting these studies, and will provide calibration samples for crucial mistag probability and background estimates.

1.2.6 The Importance of b-Tagging

In all the processes described in this section, b quarks appear in the final state. The identification (tagging) of b-quark jets will therefore play a central role in the high-p_T physics program during Run II. It can be used to reduce backgrounds from light-quark and gluon jets and to reduce combinatoric effects. If b-quark jets can be tagged already at the trigger level, the rate of triggers for these processes can be reduced without loss of efficiency. This is especially important for processes with all-hadronic final states like $t\bar{t}$ →all-jets, $WH \rightarrow q\bar{q}b\bar{b}$, $\rho_T^+ \rightarrow b\bar{c}\,b\bar{b}$, or $Z \rightarrow b\bar{b}$. Triggers for these processes are swamped by gluon-jet background. As a result, triggers with jet-p_T thresholds low enough to see a mass resonance in the 100 GeV range cannot be operated without significant prescale. Additional background rejection, like that afforded by tagging b jets, is required to exploit the full integrated luminosity delivered by the Tevatron.

This is the motivation for proposing a trigger processor for the data from the SMT at trigger level-2. This processor reconstructs tracks from charged particles precisely enough to identify tracks from the decay of long-lived b quarks. The b quarks produced in $p\bar{p}$ collisions live long enough (about 1.5 ps) to travel several millimeters. Tracks of charged particles from b quark decays can therefore be identified by their large impact parameter with respect to the beam trajectory.

1.2.7 Sharpening of Trigger Thresholds

The L2STT will give improved momentum resolution leading to sharper transverse momentum thresholds for charged tracks, jets, and electrons at level-2 and level-3. The p_T of charge particles can be measured more precisely at level-2 if the SMT hits are included in the fit (see section 3.3), leading to sharper p_T thresholds for level-2 track triggers. The transverse momentum of jets and electrons is determined from their energy measured in the calorimeter and the position of the interaction vertex. The L2STT provides a precise measurement of the position of the interaction vertex along the beam direction which can be used in level-3 to improve the transverse momentum resolution for jets and electrons.

Sharper trigger thresholds generally lead to reduced trigger rates, which results in a more efficient use of the data acquisition bandwidth of the experiment and benefits the entire physics program. Soft thresholds allow a fraction of the particles with p_T below threshold to fire the trigger. Since the p_T spectrum of particles produced in $p\bar{p}$ collisions rises steeply as the p_T decreases, this leads to an increase in the trigger rate. A sharper the threshold cuts out this background below threshold and thus reduces the trigger rate. We quantify this effect in section 3.5 using a simulation.

1.3 The DØ Trigger System

During Run II, the proton and anti-proton beams will 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 sustained rate of 50 Hz. The function of the trigger system is to recognize a beam crossing during which an interesting event occurred and then to initiate the storage of the event for off-line analysis. In order to perform this function efficiently, the DØ experiment will have a three-level trigger system in Run II.

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, the CFT, and the preshower detectors, 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 5–10 kHz. As soon as the digitization is initiated, the detector stops accepting new events.

Level-2 consists of an array 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 [15] which collects the data from all 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 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 must reduce the Level-1 accept rate of 10 kHz by an order of magnitude to about 1000 Hz. The proposed L2STT would be another element in the array of preprocessors. It interfaces with the global level-2 processor in the same way as the other preprocessors and requires no modifications to the approved level-2 architecture. After addition of the L2STT there will be preprocessors for all major detector components, thereby maximizing the performance of the level-2 system. Figure 2 shows a block diagram of the level-1 and level-2 trigger system.

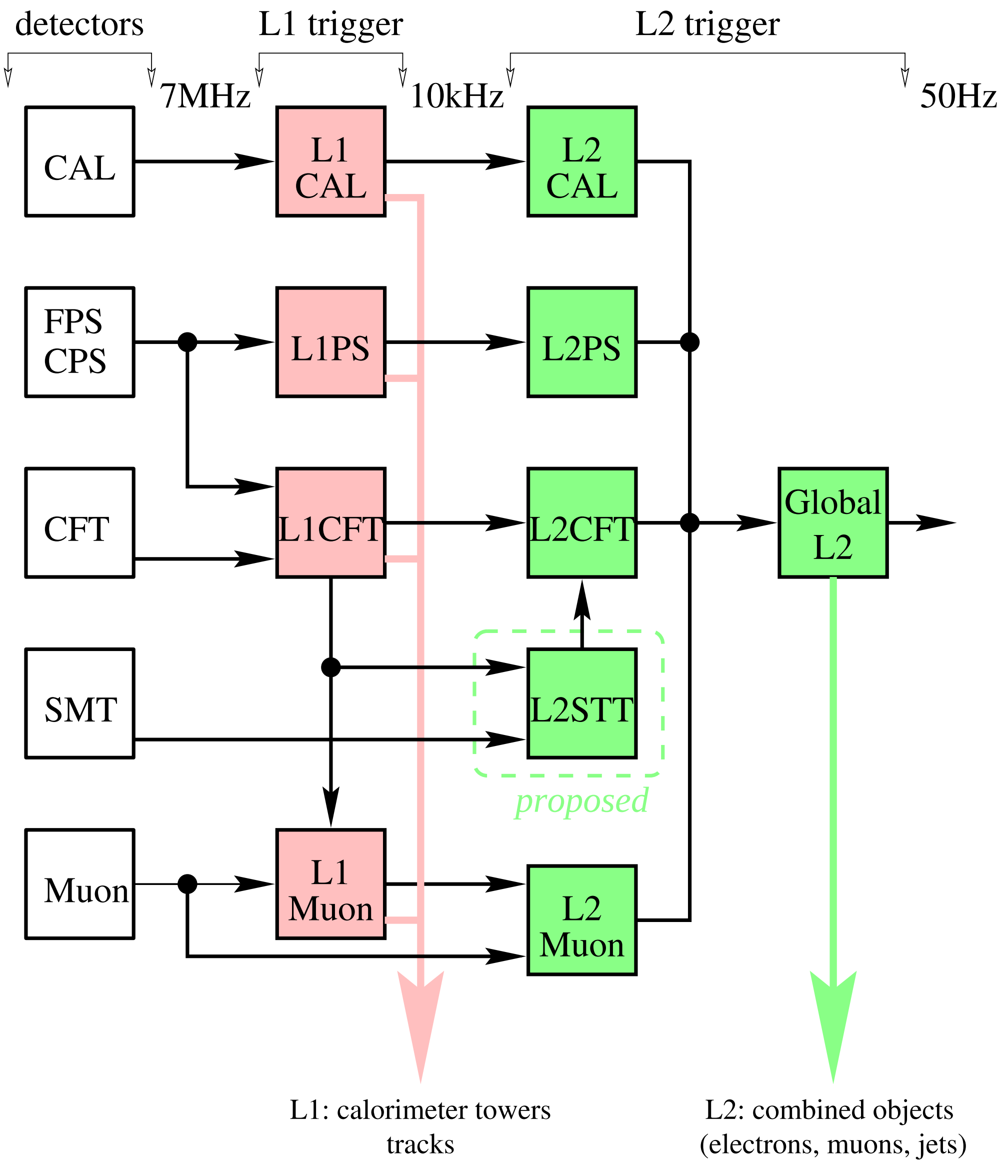

Figure 2: Block diagram of the level-1 and level-2 trigger systems.

Level-3 consists of a farm of fast high-level computers and performs a simplified reconstruction of the entire event. Events that satisfy the desired characteristics can be written to a permanent storage medium at a sustained rate of 50 Hz. The level-3 system has 25 ms to reconstruct an event. This may not be long enough to unpack the hits from the SMT and use them to find the number and z-positions of the primary interaction vertices. The L2STT solves this problem by finding the hit clusters in the SMT and the interaction vertices in level-2 and transmitting them to level-3. The high degree of parallel processing and the dedicated hardware processors of the L2STT can perform this task in much less time than the high level serial processors of level-3. Knowledge of the z-position of the primary vertex at level-3 is required for the precise calculation of jet and electron p_T from their energies measured in the calorimeter. This is provided by the L2STT.

1.4 The DØ Silicon Microstrip Tracker

Figure 3 shows a view of the SMT which consists of six 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 in beam direction, which requires the barrel section to be long (72 cm). The disks extend the acceptance of the detector to forward tracks.

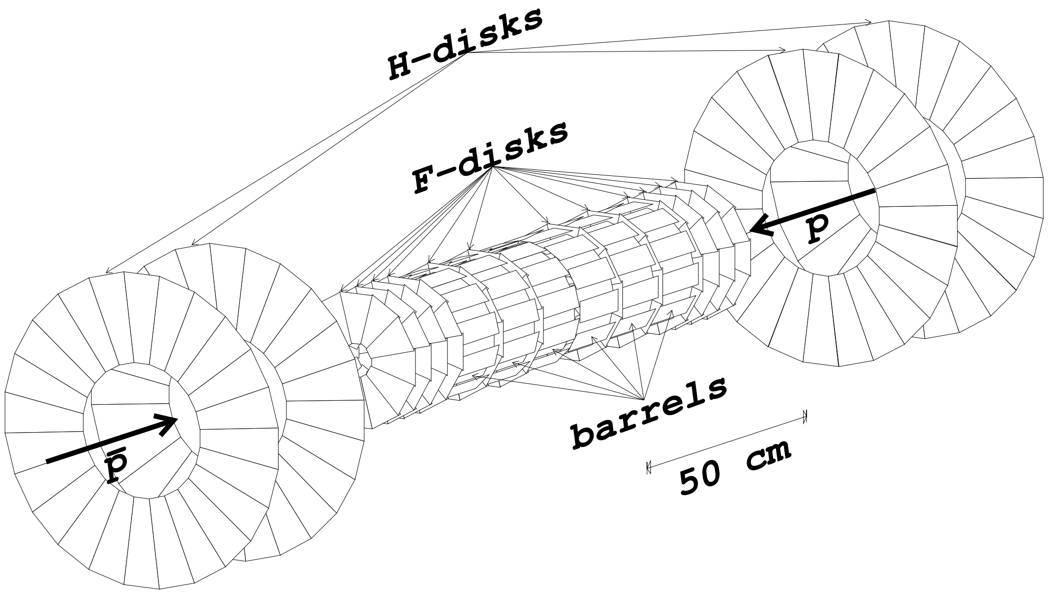

Figure 3: View of silicon microstrip tracker.

The barrel sections consist of rectangular silicon detectors arranged in four layers. Figure 4 shows the arrangement of the detectors in the plane transverse to the beam direction. Each layer consists of two overlapping sublayers, referred to as a and b. All detectors are segmented into axial strips (parallel to the beam axis). The detectors in layers 1 and 3 of the central four barrel sections also have strips that form an angle of 90° with the axial strips, so that they directly measure the z-coordinate of the hits. The z-axis points along the direction of the proton beam. The detectors in layers 2 and 4 of all six barrel sections have stereo strips that form an angle of 2° with the axial strips. Only the axial strips are used in the track fits performed by the L2STT. The hits in the 90° strips are used in the determination of the z-position of the interaction vertex.

As shown by the shaded regions in Figure 4 each barrel section can be divided into 12 sectors 30° in azimuth. Due to the azimuthal overlap of adjacent detectors, almost all tracks hit detectors

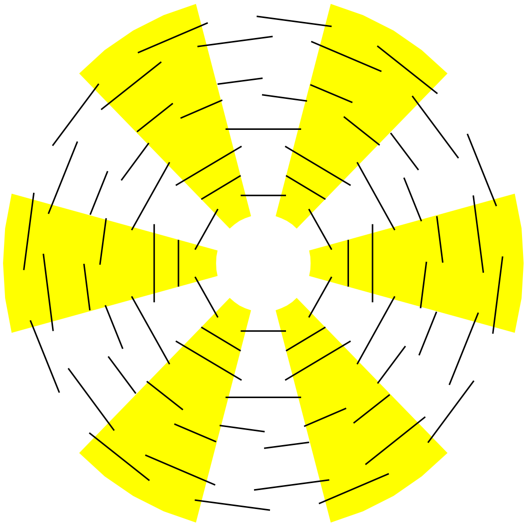

Figure 4: Arrangement of barrel detectors in transverse plane.

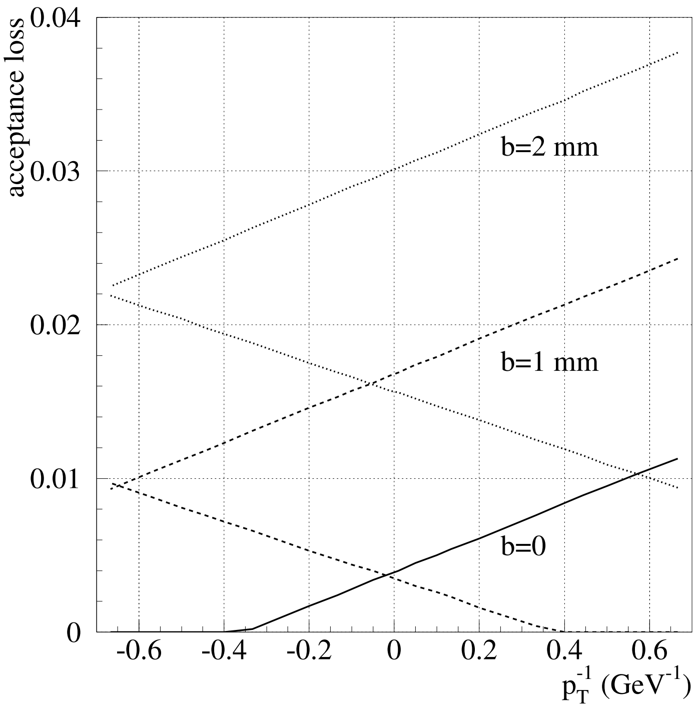

Figure 5: Acceptance loss due to tracks that cross between SMT sectors for different impact parameter values b. The two branches correspond to positive and negative values of b.

that belong to the same sector in all four layers. The acceptance loss if we miss tracks that cross sectors depends on their p_T and impact parameter b. Figure 5 plots the fraction of tracks that are not contained in a single sector versus the curvature ($=1/p_T$) of the track for several impact parameter values. The plot reflects the slight asymmetry of the detector for 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 loss in acceptance is driven by the small overlap between the active regions of the detectors in layer 3. If tracks with hits in only the other three layers (1, 2, and 4) are accepted, there is no acceptance loss at all. We can thus safely ignore tracks that cross between sectors and divide the SMT into 12 independent sectors.

The silicon detectors are read out by SVXII chips, mounted on a flex circuit (HDI) which is installed directly on the detectors. The SVXII chips each contain 128 channels of preamp, analog pipeline, and ADC. The digitized signals are read out sparsified into sequencer cards. The sequencer cards are located in the collision hall near the detector. Each sequencer card accepts input from four pairs of HDIs. The data from each pair is transmitted over a fiberoptic link to VME readout buffer (VRB) cards [16] located in the counting house. The VRB cards buffer the data for readout to the level-3 trigger and to permanent storage.

2 Conceptual Design

2.1 Overview

This section describes the conceptual design of the L2STT hardware. The L2STT 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 $r\phi$ track candidates to the L2STT. Each track candidate defines a road in the SMT. The L2STT receives all data from the SMT barrel detectors and forms hit clusters from the raw data. It filters the $r\phi$ hits, retaining only the hits that fall into one of the roads defined by the CFT data, finds the hits most consistent with the corresponding CFT track and fits a trajectory to the selected $r\phi$ hits and the CFT track. Then it communicates a list of fitted track parameters to the level-2 CFT preprocessor (L2CFT), which merges the information with the level-1 CFT information and formats the data for transmission to the global level-2 processor. The hits in the 90° strips are used to determine the z-positions of the interaction vertices, which are directly communicated to the global level-2 processor.

As discussed in section 1.4, the SMT is divided into 12 sectors, each covering 30° in azimuth. The L2STT ignores tracks that cross sectors which substantially simplifies the design. The L2STT can then be broken down into 12 subunits, each covering one sector, that do not communicate with each other. Two such subunits share a VME crate so that there are six such crates. Each crate holds:

- 1 fiber road card. It receives the information from the CFT level-1 trigger and broadcasts it across the backplane to the trigger cards.
- 9 trigger cards. They convert the CFT information into roads in the SMT and receive the SMT data from the sequencer cards, find hit clusters and sort the clusters by roads. The data processing on the trigger cards takes place in field programmable gate arrays (FPGA).
- 1 or 2 track fit cards. These cards perform the track fitting.
- 1 vertex card. This card determines the z-position of the interaction vertex.
- 1 Bit3 card for parameter download and testing purposes.
- 1 VBD (VME Buffer Driver) for readout of data to level-3.

The fiber road cards, trigger cards, and track fit cards are custom designed VME64-compatible PC boards. The electronics in the L2CFT, consisting of processors and communication links, are built from components used in other parts of the DØ trigger system. In this way the L2STT will fit seamlessly into the existing level-2 design. Figure 6 shows the crate structure of the L2STT.

2.2 Fiber Road Card

The fiber road card serves five functions in the overall L2STT design.

- receive data from the CFT level-1 trigger;
- broadcast the CFT data over a special purpose bus on the backplane to the trigger cards;
- receive control signals from trigger framework over the serial command link (SCL);
- control buffering and readout of the data to level-3;

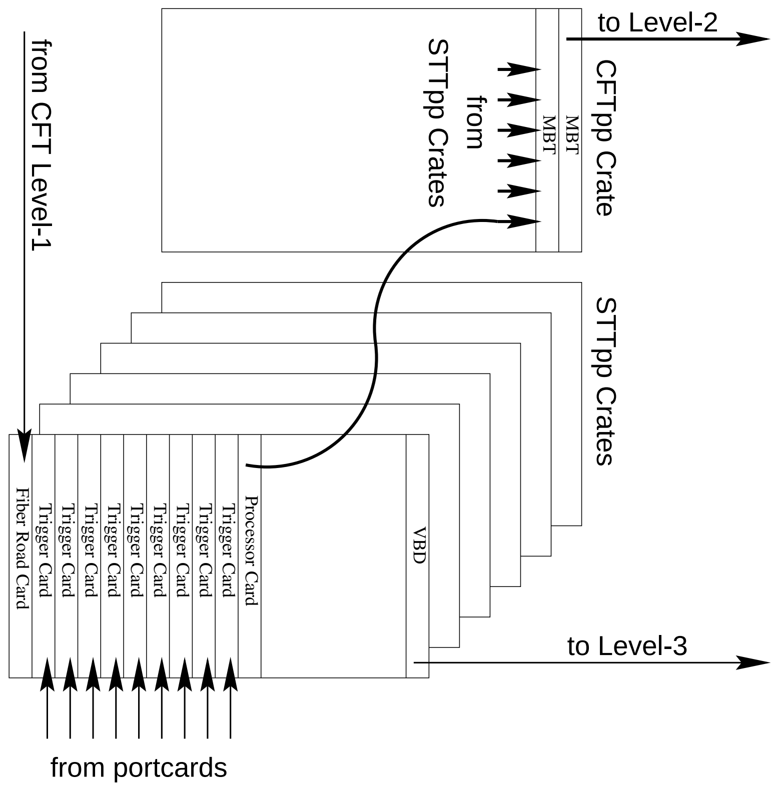

Figure 6: Crate structure of L2STT.

- VME controller.

One of these cards is required per L2STT VME crate. A block diagram of the fiber road card is given in Figure 7.

The CFT level-1 trigger transmits up to 32 CFT track candidates over fiberoptic cable to each fiber road card. For each track candidate it sends the identities of the fibers hit in the inner and outermost CFT layers. Tracks with $p_T > 1.5$ GeV and $|b| < 2$ mm that pass through a given L2STT sector can hit any one of 773 fibers in the outermost CFT layer, spanning a region of $60°+19°$ in azimuth [17]. The extra 19° account for the curvature of the track and its finite impact parameter. Tracks that pass through a given outer fiber can hit any one of 28 fibers in the inner CFT layer. Thus at least 15 bits are needed 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 L2STT crates, is currently under study.

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. The CFT information is broadcast to the trigger cards over a backplane bus.

2.3 Trigger Card

2.3.1 Overview

The trigger card fulfills the following tasks:

- receive CFT data from the fiber road card over the backplane and convert it to roads in SMT;

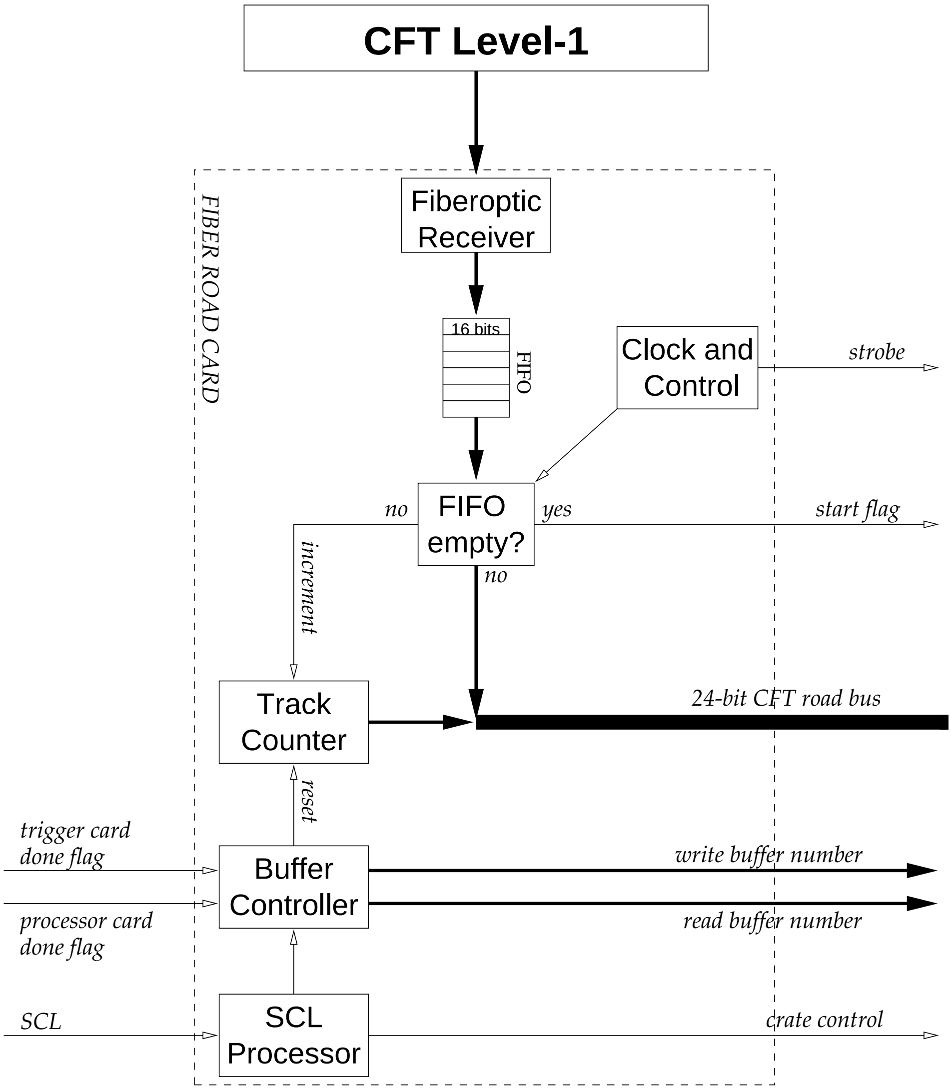

Figure 7: Block diagram of fiber road card.

- receive SMT data from the sequencers over optical fibers and find hit clusters;
- buffer SMT data;
- select $r\phi$ hits that are inside the roads defined by the CFT data;
- transmit the $r\phi$ hits in the roads to the track fit cards;
- transmit all hits from the 90° strips to the vertex cards;
- buffer hit clusters for readout to level-3.

Figure 8 shows a block diagram of the trigger card.

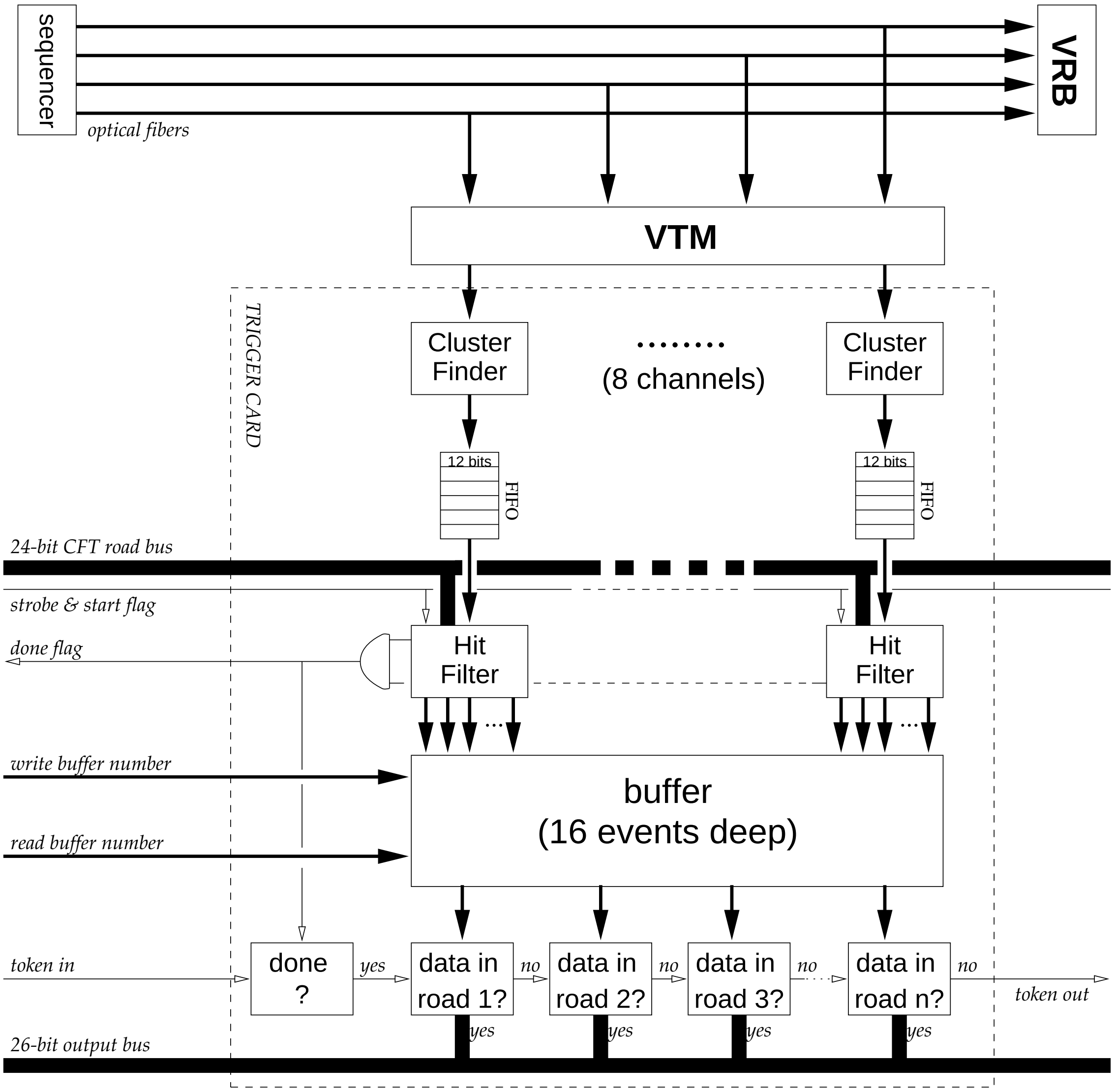

Figure 8: Block diagram of trigger card.

2.3.2 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 a track number This track number uniquely identifies each track for a given event. For a system capable of handling 200 tracks, eight bits are needed. The bus thus requires 24 signal lines and one strobe line. It can be run at 25 ns per cycle. According to Monte Carlo simulations [18], we expect on average less than two CFT tracks per crate. The maximum number of tracks expected is 25, the maximum number transmitted by the CFT level-1 is 32. It therefore takes on average 50 ns and at most 800 ns to transmit the CFT data across the backplane to the trigger cards. In addition, it takes about 2.4 μs after a level-1 trigger accept for the CFT data to arrive at the fiber road cards. That means the road addresses are latched in the trigger cards no later than 3.2 μs after a level-1 trigger is accepted.

2.3.3 Conversion of CFT Tracks to 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 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. 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 and for each track we need to look up 20 bits. For each channel, we therefore need a lookup memory of 54 kbyte. The conversion from CFT track addresses to SMT strip ranges can be performed synchronously and introduces only a small latency.

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 9 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 impact parameters $|b| < 2$ mm, the road widens as indicated by the dashed lines.

Figure 10 shows the width of the road in azimuth ($\Delta\phi$) 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 11 shows the road widths in arclength ($\Delta s = \Delta\phi \times r$). Table 1 lists the road widths at the mean radii of the SMT layers. In the table "a" and "b" refer to the two overlapping sublayers that make up one SMT layer.

2.3.4 Input from SMT Front-Ends

Optical fibers transmit the SMT data from the front-ends to the VRB cards in the counting house. The data from each silicon detector consists of alternating 8-bit address and signal words. The data from two silicon detectors are transmitted over each fiber at a rate of 106 Mbyte/s. Passive splitters in the fibers create a data path into the L2STT, where the fibers plug into the same VME transition module (VTM) as used for the VRBs. The data are loaded into a FIFO as they are received.

Digitizing the SVX-II data takes about 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

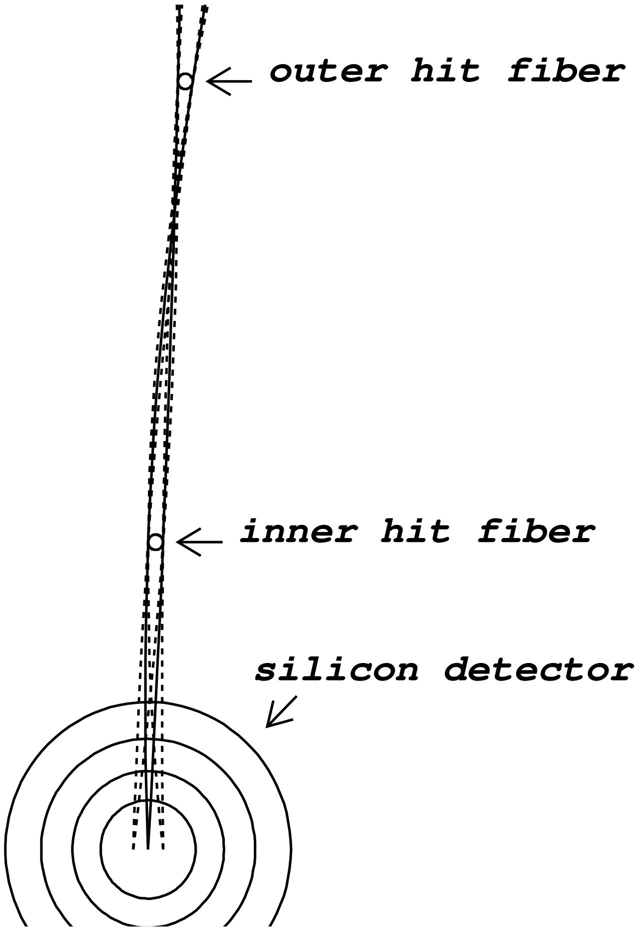

Figure 9: Road definition in the SMT.

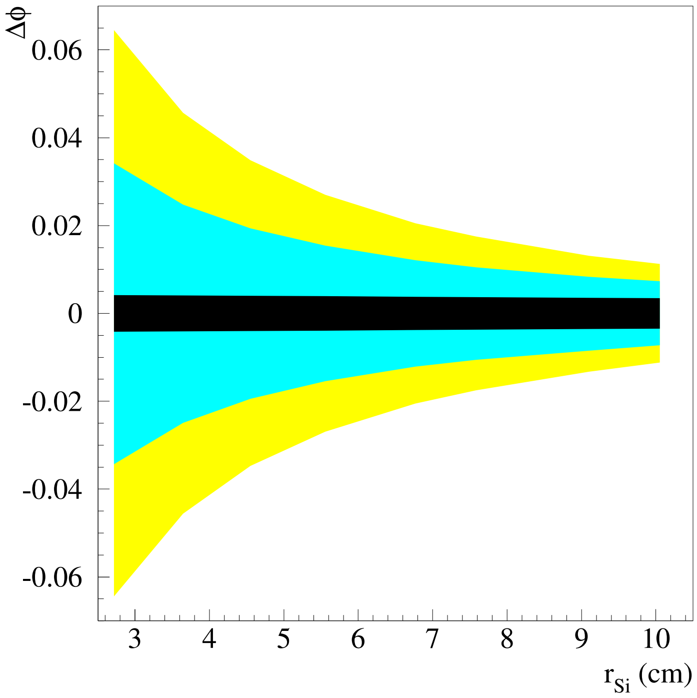

Figure 10: Width of SMT roads versus radius.

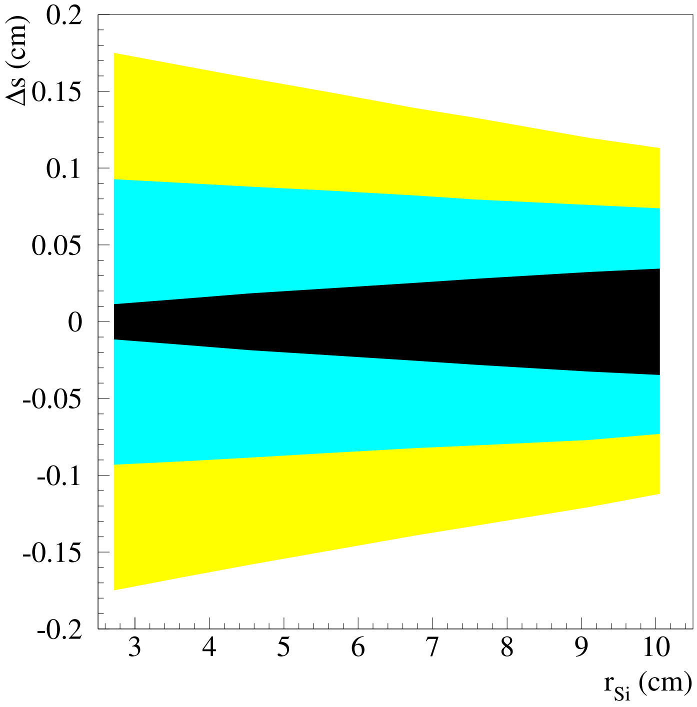

Figure 11: Width of SMT roads versus radius

Table 1: Road sizes in SMT.

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

channels. Monte Carlo simulations predict an average of 40 hits per fiber, corresponding to an occupancy of 2% [18] and a maximum of 180 hits per fiber. We add 1% occupancy from noise (see analysis of test beam data [19]). We thus have 60–200 channels to read out. At 53 MHz this will take 1.1–3.8 μs.

The L2STT has to provide seamless coverage of all barrel segments that belong to one SMT sector covering 30° in azimuth, because many tracks cross from one barrel section into another. Very few tracks are not contained in a single SMT sector (section 2) 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 L2STT crate. In the six barrel sections, there are 36 detectors per sector. Each VTM accepts four fibers and each fiber carries the data from two detectors so that each trigger card has eight identical channels, each processing the data from one detector. We can therefore accomodate two SMT sectors in an L2STT crate that is populated with nine trigger cards.

2.3.5 Hit Finding and Filtering

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 pulse heights below a threshold. It has to be determined whether these operations need to be performed on a strip-to-strip, SVXII chip-to-chip, or global scale. The FPGA then scans the corrected pulse heights for local maxima, *i.e.* strips with larger pulse heights 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 pulse height-averaged center of gravity of up to five adjacent strips, centered on a seed strip [20]. The hit finding runs synchronously with the incoming data so that it introduces only a small latency ($\ll 1$ μs). For each cluster the FPGA outputs 13 bits: the seven bits used to identify the strip number plus an additional two bits of precision added by the cluster finding algorithm, three bits to identify the SVXII chip, and one bit to indicate whether the cluster is on the axial or the stereo side (see figure 12).

Based on bit 13 the clusters are stored in two FIFOs. Only clusters from the axial strips participate in the track finding. The clusters from the 90° stereo strips are used to find the primary vertex (section 2.7). All clusters are buffered for readout to level-3.

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 12 are compared to the road limits.

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

Figure 12: Bit allocation in output of cluster finder.

Since there can be up to 32 roads, 64 10-bit comparators are required. This hit filtering process is performed as the SMT data arrive and will keep up with the data so that does not cause a delay but only a small latency equal to the time it takes to process one cluster ($\ll 1\ \mu$s). Figure 13 shows a block diagram of the hit filter FPGA.

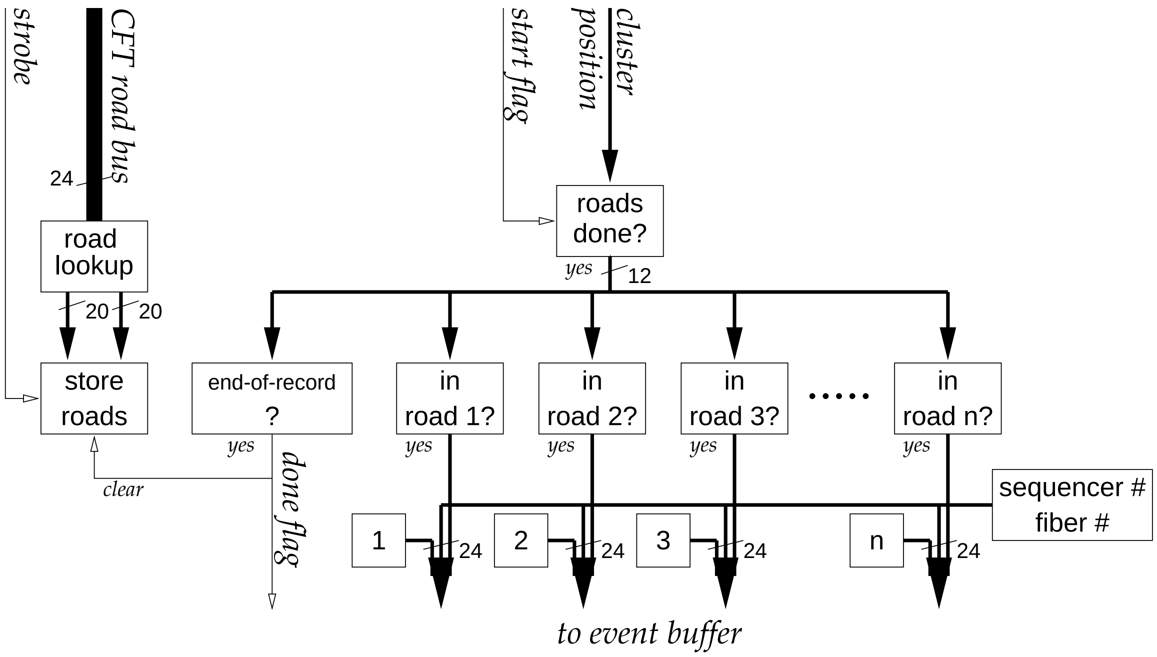

Figure 13: Block diagram of hit filter FPGA.

2.3.6 Event Buffering

The output of the hit finder FPGA is stored in a buffer 16 events deep. When all data from an event have been processed a "done" flag is set. The logical and of the flags from the eight channels 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. Adding up the time delays and latencies quoted for each step we find that this process is expected to be completed about 5–8 μs after a level-1 trigger is accepted. At this time the L2STT is ready to accept another event.

2.3.7 Transfer of $r\phi$ Hits to Track Fit Card

For each hit 26 bits need to be transferred: 7+2 bits for the fractional strip number, 3 bits to identify the SVXII chip within a fiber, 2 bits to specify the fiber within a sequencer, 4 bits to identify the sequencer, and an 8-bit track number (figure 14). The track number is generated in the fiber road card as the tracks are received. The lower 12 bits are the same as the output of the cluster finder (figure 12).

26 ... 19	18 ... 15	14 13	12 ... 10	9 1
track	sequencer	fiber	chip	strip

Figure 14: Bit allocation in output to track fit card.

The data transfer occurs over a dedicated backplane bus which provides unidirectional, data-driven broadcast from the trigger cards to the track fit card. It 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 the track fit 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 track fit card.

There are on average 42 and at most about 500 clusters in all roads of an L2STT crate [18]. If the bus operates at 25 ns/cycle, the transmission of these clusters and the CFT data to the track fit card will take on average 1.1 μs and at most 13.3 μs.

2.3.8 Transfer of z Hits to Vertex Card

The hits in the 90° strips of the silicon detectors in layer 1 and 3 of the central four barrel sections give information about the z-position of charged particle hits. We therefore call them z hits. They can be used to determine the z-position of the interaction vertices in an event. For this purpose they are transferred to vertex cards over a backplane bus. The transfer protocol is similar to that used to transfer the $r\phi$ hits to the track fit card. For each z hit only 16 bits need to be transferred: 7 bits for the strip number, 3 bits to identify the SVXII chip within a fiber, 2 bits to specify the fiber within a sequencer, and 4 bits to identify the sequencer (figure 15).

16 ... 13	12 11	10 ... 8	7 1
sequencer	fiber	chip	strip

Figure 15: Bit allocation in output to vertex card.

The expected hit occupancy for 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 2. We thus expect some 100 z hit clusters in one L2STT crate. If the bus operates at 25 ns/cycle, the transmission of these clusters to the vertex card will take on average 2.5 μs.

Table 2: Average number of z hit clusters in 30° sectors.

Layer	1 interaction	3 interactions	6 interactions
1	6	13	26
3	4	10	17

2.4 Track Fit Card

2.4.1 Overview

The Track Fit Card performs three basic functions:

- receive $r\phi$ hits in SMT and CFT track information from the trigger cards and convert from hardware address to physical coordinates;
- perform the track fit;
- communicate the fit results to the CFTpp.

2.4.2 Input Stage

The input stage consists of bus receivers, simple control logic, a table to map roads to processors, and a look-up table to provide conversion from hardware addresses (detector, SVXII chip, and strip numbers) to positions in (r, ϕ) coordinates.

Each CFT road is associated with a unique CFT track number and a look-up table associates the track number with a processor. The table is reset at the start of each event and built "on the fly" as data for each track are received. This logic requires very modest resources: 32 8-bit comparators, 32 8-bit registers, and minimal control logic.

A second table is used to convert the hardware addresses to physical coordinates. The simplest possible approach has been adopted. The look-up table contains one entry for every possible cluster position in the SMT sector. This implies between 128k and 256k entries in the look up table for each 30° sector. Each entry contains the corresponding (r, ϕ) coordinates. If the radius is stored in units of 10 μm and the azimuthal angle granularity is derived from the smallest angular resolution resulting from the 1/4 strip precision of the cluster address, the table needs at most 32 bits for each entry, 14 bits for r and 18 bits for ϕ.

Although the coordinate conversion table is quite large, the download is not a significant problem because the table data change only when the detector alignment changes. The detector is designed to require infrequent relignment, and RAM with battery backup could be used to eliminate the need to download after power failures.

2.4.3 Processing Stage

During the data loading phase, the input stage is used to fill the internal memory of the processors with the initial track parameters and positions of associated silicon hits. Then the processors begin fitting a track parametrization to the hits. The track parameters are the signed impact parameter b, the azimuth of the track at the point of closest approach to the beam ϕ_0, and the curvature κ.

The number and type of processor required depends on the complexity of the track reconstruction algorithm to be performed. The final choice will be based on cost and performance. Simulation studies of algorithm and processor options are in progress (section 3.2).

2.4.4 Output Stage

The output stage is used to transfer the fit results to the CFTpp. The transmitted information consists of the CFT track information and number (24 bits), 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 be over a fast serial link running at 212 Mb/s to a Magic Bus Transceiver (MBT) Card [21] in the CFTpp crate. The MBT cards are part of the approved architecture of the level-2 trigger system and their design is in progress.

An upper bound on the data transfer time can be derived under the assumption of 32 bits of precision for each of the four quantities from the track fit so that there are 19 bytes per track. However, the output can probably be packed to about 8 bytes per track. The bit count is summarized in Table 3 To transfer the maximum of 32 tracks that each L2STT crate can process per event takes about 10 μs.

Table 3: Data transmitted to the CFTpp crate for each L2STT track.

Quantity	number of bits		
	full	packed	(precision)
CFT hits and track number	24	24	n/a
b	32	15	(1 μm)
p_T or κ	32	11	(0.1 GeV)
ϕ_0	32	8	(0.025 rad)
χ^2	32	6	(0.2)
total	152	64	

2.5 Communication with Global Level-2 Processor

The results of the track fits performed on the processor cards in all six L2STT crates is transferred to the CFTpp which acts as a concentrator to consolidate the six links into a single stream. The CFTpp is a part of the approved level-2 architecture and has the standard DØ level-2 preprocessor structure, consisting of

MBT cards to receive data from the CFT level-1 trigger and the L2STT, and to send processed track information to the global level-2 processor via a fast serial link.

DEC-Alpha Processors to match CFT and L2STT information, sort and format the data for transmission to the global level-2 processor.

2.6 Communication with Level-3

For diagnostic and monitor purposes, all information created by a processor must be read out by level-3 on every event. The L2STT creates two kinds of data: hit clusters and 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.

To save unpacking time in level-3 it may be desirable to read all hit clusters in the SMT out for further processing in level-3. These data can be read out from the trigger cards. 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 L2STT crate. Every cluster position requires 18 bits (see figure 14) and all clusters correspond to 0.6–2.3 kbytes 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. Since the complete track information is transferred to the CFTpp it can be read out to level-3 from there such that there is no need to read out the track fit cards to level-3.

2.7 Vertex Processor

2.7.1 Overview

The vertex card receives the z hits in the SMT from the trigger card, determines the number and z positions of interaction vertices in the event, and transmits the results to the global level-2 processor.

In contrast to the $r\phi$ tracking which can be performed locally with information from one SMT sector only, the vertex reconstruction is a global problem that can only be solved by using information from the entire detector. The biggest challenge in implementing an algorithm to find the interaction vertices is thus to bring the information from all z hits in the detector which is scattered over all 54 trigger cards in the system together on a single board.

2.7.2 Algorithm

In order to satisfy the rigid time constraints imposed by the trigger system, our algorithm reconstructs the vertices without the intermediate step of explicitly reconstructing all particle trajectories.

Figure 16 shows a side view of a simulated Tevatron event. The smooth lines originating from three points in the center of the detector are trajectories of charged particles. The three "sources" of particles are the vertices of the three proton-antiproton interactions in the event. The curvature of the trajectories are due to the magnetic field in the detector. The thick vertical lines indicate the position of annular disks of silicon microstrip detectors and the four horizontal lines above and below the intersections define the locations of the four cylindrical barrel layers of silicon microstrip detectors. Layers 1 and 3 have strips that run at 90° to the beam direction and thus measure the z position of charged particle hits.

Figure 17 shows a schematic representation of the hits in one SMT sector in the side view. The short vertical lines indicate hits registered in layers 1 and 3. Consider any one hit in layer 3. We assume it originates from a charged particle that also left a hit in layer 1. The layer 3 hit in conjunction with all the n hits in layer 1 then constrains the position of the interaction vertex to one of n discrete positions, as shown by the dashed arrows. This set of allowed vertex positions is given by the intersections with the beam of all lines defined by the hit in layer 3 paired with each hit in layer 1. We have neglected the curvature of the tracks due to the magnetic field in the center of the detector. This is quite a good approximation for tracks with a momentum component transverse to the beam direction above 1.5 GeV. We now evaluate the analogous constraints given by all hits in layer 3 and accumulate the allowed vertex positions in a histogram. The patterns of allowed vertex positions for two different hits in layer 3 differ only by an offset

$$\Delta = \frac{r_1}{r_3 - r_1}\Delta z, \tag{1}$$

where Δz is the difference of the z-coordiantes of the two hits in layer 3, r_1 is the radial position of layer 1, and r_3 the radial position of layer 3. Since the tracks point radially outward from a

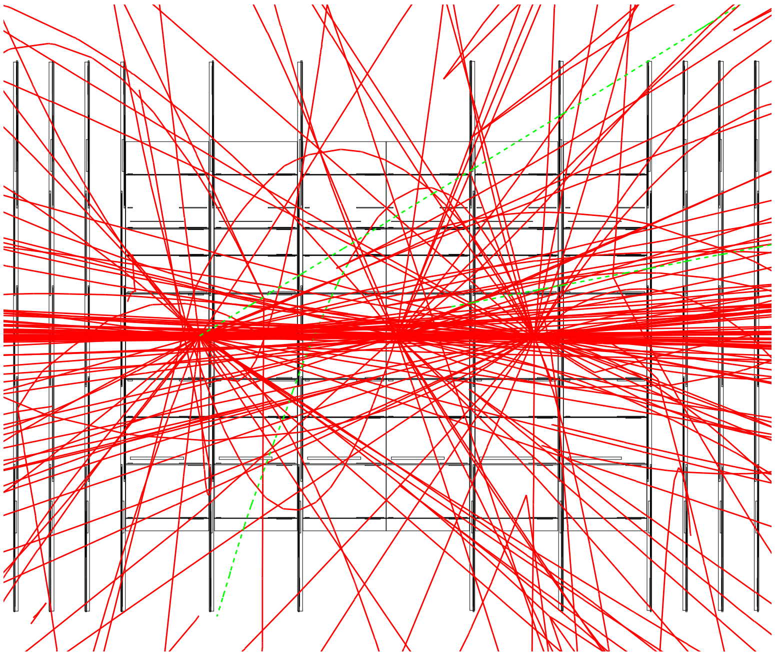

Figure 16: Sideview of a simulated Tevatron event with three $p\bar{p}$ interactions.

very small region in the center of the detector, we perform this operation for all 12 30°-sectors independently to reduce the combinatoric background that arises from unrelated hit pairs in layers 1 and 3.

Finally, we add the histograms from each of the 12 sectors to produce a histogram for the entire event. At the position of the interaction vertices there will be an excess of entries above the combinatoric background. Figure 18 displays the accumulated histogram for the event shown in figure 16. The excesses at the vertex locations are evident. A simple peak-finding algorithm returns number and positions of these vertices.

2.7.3 Implementation

On the vertex boards in each L2STT crate the z hits from each 30° sector are processed independently. The hits in layer 1 are histogrammed with 512 bins and a bin size of 250 μm. This granularity is fine enough to yield a resolution for the z-position of the vertices of about 550 μm [22]. The hits in layer 3 are loaded into a FIFO. Then a hit z_3 is taken from the FIFO in each processing cycle and the offset $\Delta = z_3 r_1/(r_3 - r_1)$ is calculated. The bit pattern defined by the inner hits is added to a histogram with this offset. This cycle is repeated until all hits in layer 3 are processed and the FIFO is empty. On average we expect 20 hits in layer 3 per sector. If a processing cycle takes 100 ns the entire process requires 2 μs. Then the histograms from the two sectors are added together. This process will be implemented in FPGAs.

The resulting histograms, one for each sector, are then added together to one single histogram, corresponding to the event histogram shown in figure 18. For this the histogram data have to all be brought together in one physical card. We transfer the histograms from five of the vertex cards to the sixth vertex card via a fast serial links running at 212 Mb/s. Assuming 4 bits per histogram bin each histogram amounts to 2048 bits and can be transferred in about 10 μs.

The last processing step consists of loading the event histogram into a DSP to identify the number and location of statistically significant peaks. Each of these peaks corresponds to an

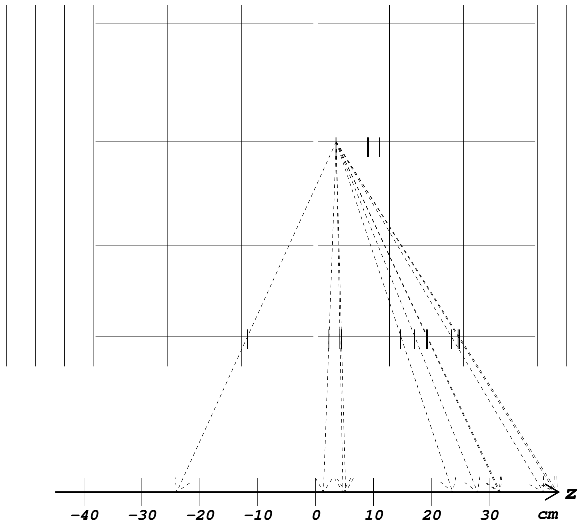

Figure 17: Constraint on z position of interaction vertex imposed by one hit in layer 3 in conjunction with all hits in layer 1.

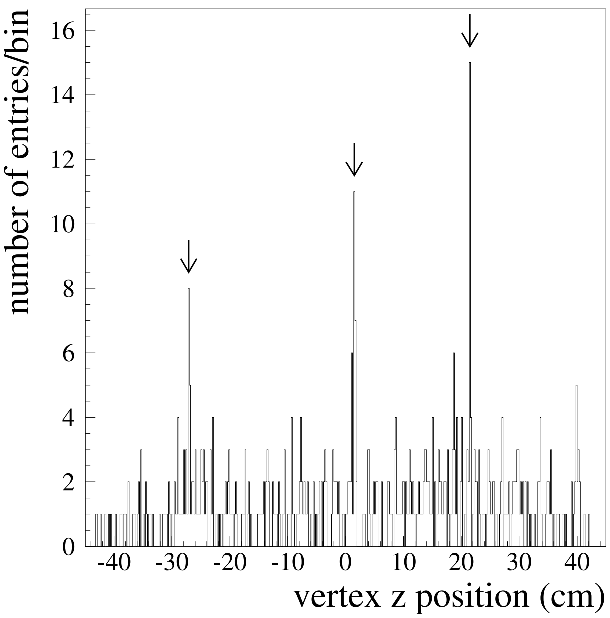

Figure 18: Histogram of possible vertex positions. The arrows indicate the true z positions of the three interaction vertices.

interaction vertex. The optimal algorithm still has to be identified, but even a simple algorithm, that accepts all bins above a threshold as vertices, works quite well. Processing time at this stage is likely dominated by the time it takes to fetch the histogram from external memory into the DSP and to put the results into an external memory. The histogram amounts to about 0.5 kB. It takes 5 clock cycles (25 ns) to read/write a word (2 bytes), resulting in a processing time of 6.3 μs.

Finally, the list of vertex positions is transferred to the global level-2 processor via another serial link. The transferred record is small, consisting of a 16-bit event tag, an 8-bit integer specifying the number of vertices and a 8-bit positions for all vertices. The entire vertex reconstruction process from the time the hit clusters are available in the trigger card takes about 25 μs, well within the time budget (section 3.1).

3 Performance

3.1 Queueing Simulations

We have used the RESQ queueing software package from IBM to build a simple model of the event flow in the level-2 trigger [23]. The model input is a random time sequence of events accepted by the level-1 trigger with a mean rate of 10 kHz. Each level-2 preprocessor is represented in the model with their processing time distribution. We adjust the parameters in the model without the L2STT so that the system operates with 5% deadtime, *i.e.* 5% of the level-1 accepts are lost because the system is busy processing previous events.

When we introduce the L2STT we assign time distributions to its processing steps. With the exception of the SVXII digitization time and the track reconstruction time, all processing times lead largely to latency in the system. We thus vary the time taken to perform the track reconstruction to determine the maximum time allowed for this step without increasing the deadtime of the system. Most time distributions are represented by double exponentials with different mean times to simulate long tails. The SVXII digitization time is fixed at 3 μs and increases the deadtime to 8%. Table 4 lists the mean processing times assumed. Some of them are larger than expected and quoted in other parts of the document, leading to a conservative estimate. According to the model, the track reconstruction can take up to 50 μs before significant additional deadtime is incurred. Figure 19 shows the deadtime incurred as a function of the mean processing time for track reconstruction.

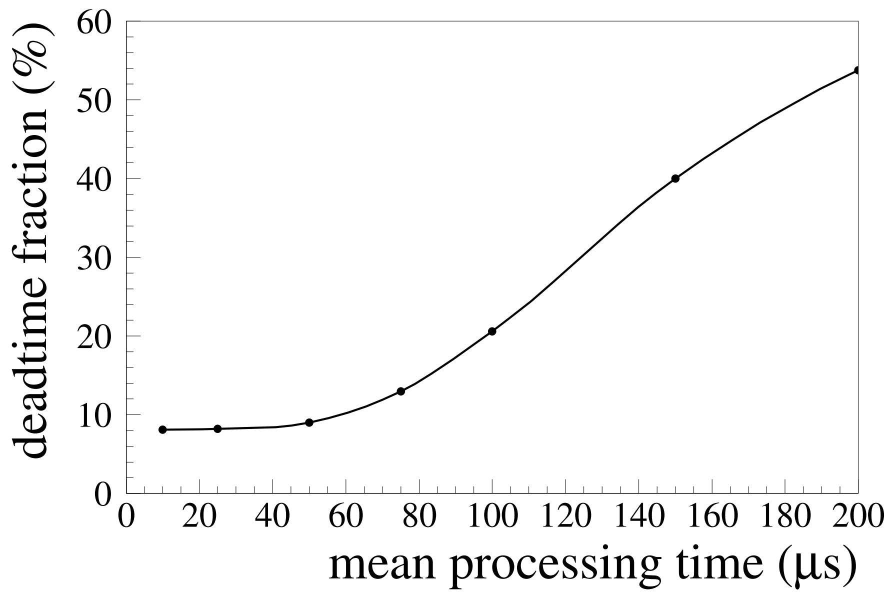

Figure 19: Deadtime incurred as a function of processing time for track reconstruction in the L2STT.

3.2 Track Reconstruction Algorithms

The track reconstruction algorithm will be implemented in the processors on the track fit cards. It has to determine the track parametrization that best fits the list of hits in a given road.

Table 4: Mean processing times for L2STT queueing model.

Processing Step	mean time	distribution
SVXII Digitization	3 μs	fixed
Data Input into L2STT	7 μs	single exponential
Hit Finding/Filtering	2 μs	double exponential
Data Transfer to Track Fit Card	8 μs	double exponential
Track Reconstruction	variable	double exponential
Data Transfer to Global Level-2	7 μs	double exponential

Once the set of hits associated with a given CFT road has been found by the trigger card, the next step in the processing is determining the track parameters in the $r\phi$ plane: the impact parameter b, the track direction ϕ_0 at the point of closest approach and the track curvature κ. There are two parts to this task: (1) determining which of the hits belong on the track and (2) determining the track parameters. In practice, the order and separability of these two steps depends critically on the algorithm being used. This section contains a description of the track definition, a description of algorithms being tested, some comparisons between the algorithms and a brief summary of initial timing studies.

3.2.1 Track Parameter Derivation

The track parameters are the impact parameter b, the track direction at the point closest to the beam ϕ_0, and the track curvature κ. We assume

- no energy loss or multiple scattering from interactions in the beam pipe or the tracking detectors, so that the trajectory is a circle in the transverse plane;
- $b\kappa \ll 1$ and $\phi - \phi_0 \ll 1$ for a silicon hit at the position (r, ϕ), so that the trajectory can be approximated by a linear function of the parameters.

These requirements are valid for all tracks with $b \leq 2$ mm and $p_T < 1.5$ GeV[1]. The expression for the track trajectory is:

$$\phi(r) = b/r + \kappa r + \phi_0. \tag{2}$$

Here r is the radial position of a point on the track and ϕ is its azimuth. The best values for b, κ and ϕ_0 are determined by minimizing the χ^2 given by

$$\chi^2 = \sum_{i=hits} \left[\frac{\phi_i - (b/r_i + \kappa r_i + \phi_0)}{\sigma_i} \right]^2 . \tag{3}$$

The set of hits consists of one hit inside the road per SMT layer and the the two CFT hits that define the road. The parameters that minimize χ^2 can be determined analytically because the fitting function is linear in the parameters.

[1]The condition $b\kappa \ll 1$ is valid for a much larger range. However, the condition $\phi - \phi_0 \ll 1$ breaks down as impact parameters approach 2 mm.

3.2.2 Hit Association Algorithms

In order to obtain the best possible performance, it is important to use only the silicon hits that belong to the charged particle track. If there is more than one hit per layer in the road we have to determine the "correct" hit. The number of hits per layer per road depends on the event topology, the luminosity and the noise in the detector. Table 5 shows the average number of hits/layer/road and the fraction of layers having two or more hits for different physical processes and luminosity conditions. Given the large number of layers with multiple hits, attention must be paid to choosing

Table 5: Hit multiplicity per road in the layer with the highest occupancy (N) and fraction of tracks with two or more hits per road in the layer with the highest occupancy (F_2) for different numbers of interactions/beam crossing. For all entries, the statistical error is negligible.

Process	# of Interactions	N	F_2
$Z \to b\bar{b}$	1	3.3	0.918
	3	4.5	0.982
$WHqb$	1	3.2	0.894
	3	5.2	0.978

hits. We are considering a number of different algorithms. These are

1. **Static Road-Center**: The hits closest to the center of the road defined by the CFT hits and the beam position are used in each layer.

2. **Dynamic Road-Center**: The hits closest to the center of the road defined by the CFT hits and the hits in SMT layer 4 are used in each layer. This algorithm requires looping over all hits in SMT layer 4.

3. **All Combinations**: All combinations of hits are considered and the one with the best fit is chosen.

4. **Best Combination at Layer**: Moving from the outer SMT layer inwards the fit is performed at the current layer using the best result from performing fits on all combinations in the preceeding layer.

In all cases the algorithm can be run in two forms. The first requires hits in all four silicon layers. The second requires hits in only three layers. Algorithm 4 is based on the *all combinations* algorithm, but the number of trial combinations is limited in a controlled manner. Fig. 20 shows the number of track candidates per road for the $Z \to b\bar{b}$(3 interactions) sample for each of the algorithms. The "Static road center" algorithm uses only one hit/layer and thus has only one track per road. The "all combinations" algorithm has the number of candidate tracks equal to the product of the number of hits in each layer, reduced slightly by the requirement that all tracks must be contained within two adjacent barrels.

We judge the performance of each algorithm using two figures of merit determined from Monte Carlo event samples: (1) the fraction of reconstructed tracks having identically one Monte Carlo track contributing hits (F_{MC}) and (2) the χ^2 of the fitted parameters with respect to the Monte Carlo track that contributes the largest number of hits. Table 6 presents the averages of these comparisons for the samples listed in table 5.

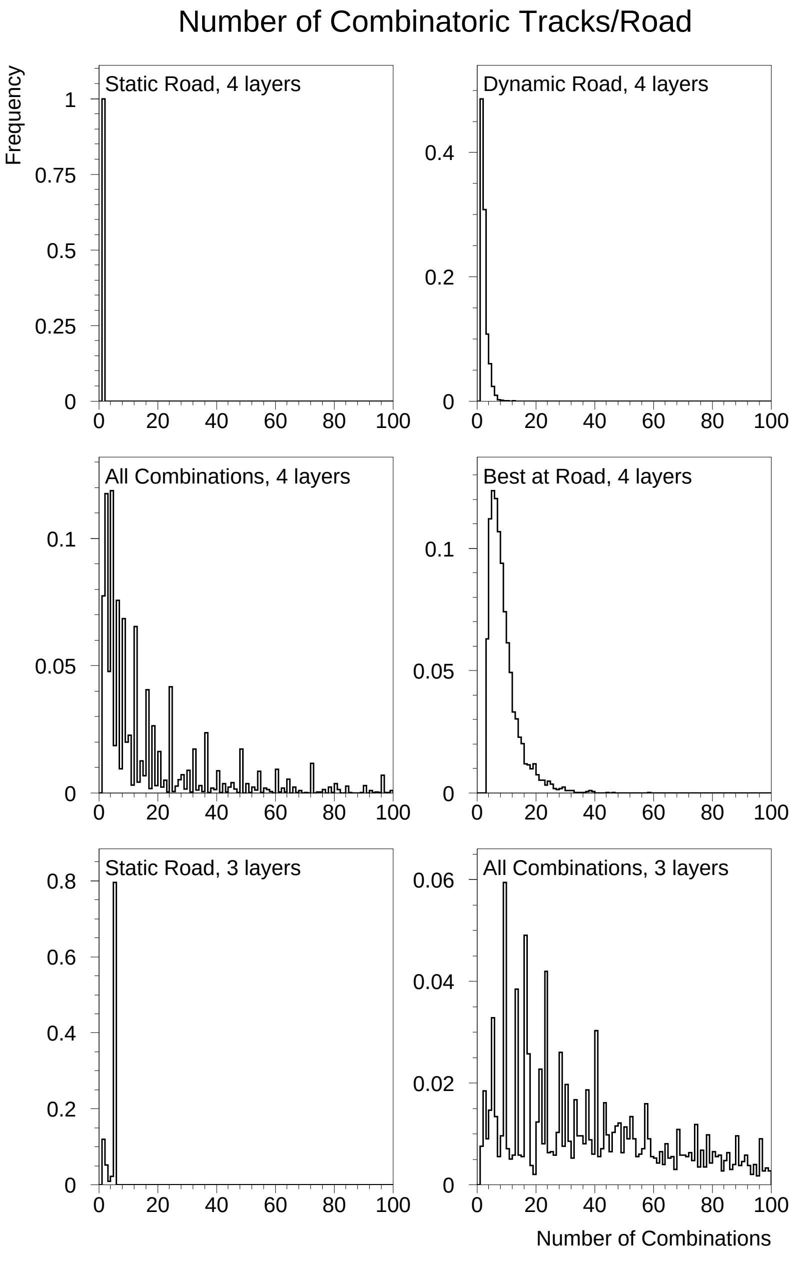

Figure 20: Number of trial tracks per CFT road for each of the algorithms. The data sample is simulated $Z \to b\bar{b}$(3 interactions).

Table 6: Tracking finding performance.

Fit Algorithm/ Minimum Hit Layers	Variable	$Z \to b\bar{b}$ 1 int.	$Z \to b\bar{b}$ 3 int.	$WH \to q\overline{q'}b\bar{b}$ 1 int.	$WH \to q\overline{q'}b\bar{b}$ 3 int.
Static Road/4	F_{MC}	0.53	0.43	0.58	0.36
	χ^2	1.4	1.5	1.4	1.1
Static Road/3	F_{MC}	0.76	0.69	0.80	0.54
	χ^2	0.7	0.8	0.7	1.0
Dynamic Road/4	F_{MC}	0.41	0.28	0.46	0.18
	χ^2	1.4	1.5	1.4	1.1
All Combinations/4	F_{MC}	0.65	0.60	0.68	0.52
	χ^2	1.2	1.3	1.2	0.8
All Combinations/3	F_{MC}	0.79	0.71	0.80	0.44
	χ^2	0.4	0.5	0.4	0.9
Best at layer/4	F_{MC}	0.51	0.43	0.54	0.40
	χ^2	1.3	1.5	1.3	0.6

The simulations described in the physics section use only algorithm (1). The comparisons in table 6 indicate that algorithm (3) or its variant have somewhat higher purity than (1) at high luminosity. To make the the final processor choice it is important to understand the gains these algorithms offer, because they require significantly more processing power. Trigger rate studies using the more complex algorithms are in progress.

3.2.3 Timing Studies

A number of different processor options are being considered. The leading choices are (1) a Digital AXP-based multi-processor board, (2) 1 GFlop TMS320C6701 floating point DSPs, or (3) arrays of programmable logic devices. The final choice will be driven by the time constraints imposed by the fitting algorithm which gives the necessary physics performance.

The samples in table 6 above have been timed on a 500 MHz AXP processor. The results are given in Table 7 on a per track basis. Fig. 21 shows the timing distributions for the $Z \to b\bar{b}$(3 interactions) sample for each of the fitting algorithms. The drastic effect of combinatorics producing tails is clearly seen. Thus, for a maximum of 32 tracks per section, the AXP solution works for

Table 7: Execution times/track (μs) for each algorithm.

Fit Algorithm/ Minimum Hit Layers	$Z \to b\bar{b}$ 1 int.	$Z \to b\bar{b}$ 3 int.	$WH \to q\overline{q'}b\bar{b}$ 1 int.	$WH \to q\overline{q'}b\bar{b}$ 3 int.
Static Road/4	3.8	4.1	3.8	3.9
Static Road/3	12.1	12.5	12.3	12.6
Dynamic Road/4	7.1	8.2	7.0	8.2
All Combinations/4	16.7	18.7	16.2	19.9
All Combinations/3	33.4	38.3	29.7	33.5
Best at layer/4	10.5	11.2	10.6	11.8

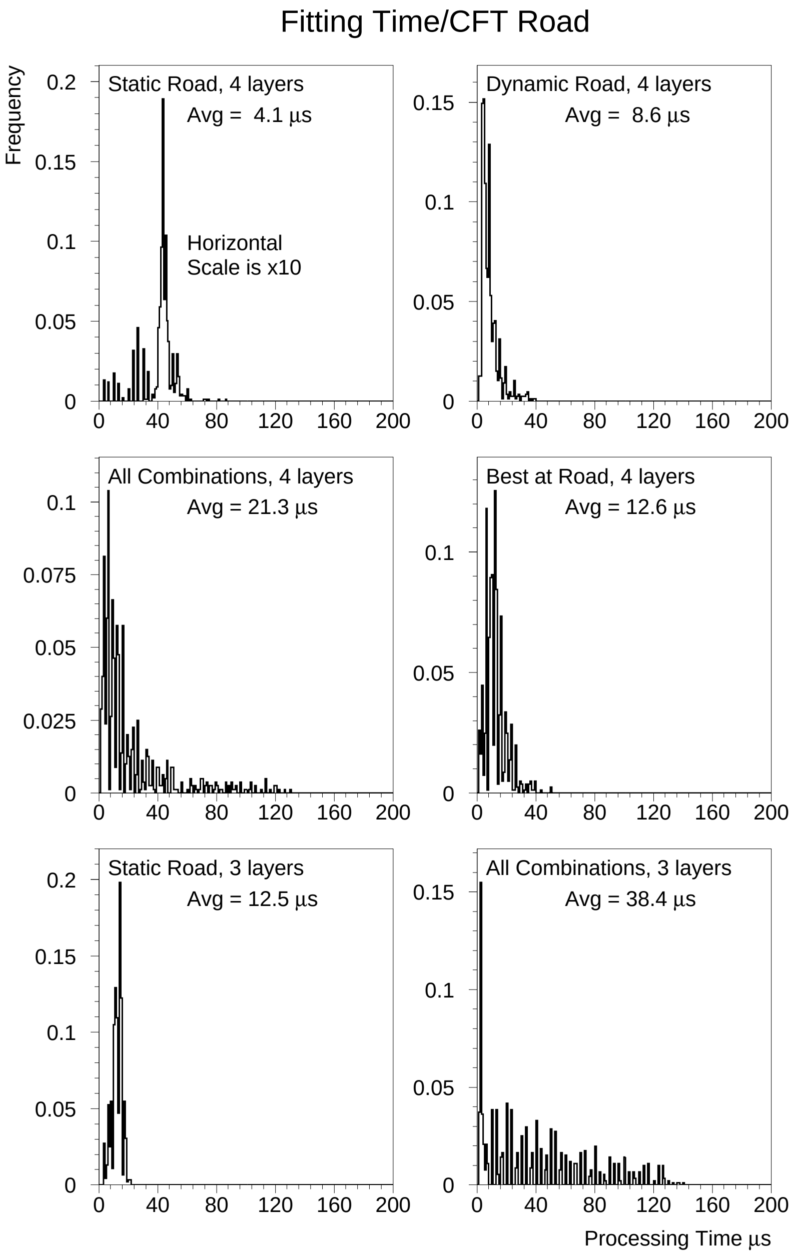

Figure 21: Fitting time per CFT road for each of the algorithms. The data sample is simulated $Z \to b\bar{b}$(3 interactions).

the 4-layer form of algorithm (1) provided there are roughly four processors per sector. The time budget cannot be met for algorithm (3) unless a sector has 32 AXP processors, one per track. The need to quantify the physics performance of the various algorithms becomes clear here.

Studies for a TI320C6x-based DSP solution indicate that a fixed-point DSP cannot provide the necessary performance because the floating point calculations needed to accomodate the dynamic range are emulated in a fixed point processor. Initial estimates indicate that algorithm (1) could be performed in roughly 5 μs on the corresponding floating point DSP. Further studies are being conducted for DSP- and FPGA-based solutions. The FPGA solution is attractive because a high degree of parallel processing can be achieved. This could allow algorithm (3) or its variants to be performed in nearly constant time. The key issue is the speed of floating point multipliers and adders implemented in programmable logic. The preliminary design and cost estimate presented in this document are based on DSPs with one processor/track road.

3.3 Expected Resolutions

3.3.1 Transverse Momentum

We use Monte Carlo simulations of single electrons and muons with $1.5 \leq p_T \leq 15$ GeV in the pseudorapidity range $|\eta| \leq 1.0$ to measure the resolution achieved by the track reconstruction in the L2STT [24]. We simulated the response of the detector to these particles using the GEANT program [25]. The track reconstruction algorithm for the L2STT at level-2 used in these simulations corresponds to algorithm (1) in section 3.2 with the following exception: if reduced $\chi^2 > 4$, the SMT hit with the largest contribution to χ^2 is dropped and the remaining hits were fit again. All tracks with hits in at least three SMT layers and reduced $\chi^2 \leq 4$ were considered "good" tracks.

We measure p_T in three different ways:

1. "L2STT": using the CFT hits in the inner and outer layers and the SMT hits (this corresponds to the measurement performed by the L2STT).

2. "CFT(A)": using the CFT hits in the inner and outer layers only (this is the information reported to the level-2 trigger and thus corresponds to the measurement of p_T that is available in the level-2 trigger without the L2STT);

3. "CFT(B)": using all eight CFT hits (this corresponds to the p_T measurement that the CFT level-1 trigger performs internally;

Figure 22 shows that the L2STT improves the p_T resolution in the level-2 trigger by a factor 2–3, depending on p_T.

The improved p_T resolution allows the use of an E_T/p_T cut for electron identification at level-2. Fig. 23 shows a comparison of E_T/p_T with p_T calculated as in CFT-L2 and as in L2STT for electron tracks with $p_T = 5$ and 20 GeV. The better resolution results in an improvement in the background rejection for the E_T/p_T cut at level-2 by a factor of 2 for $p_T = 20$ GeV and by a factor 4 for $p_T = 5$ GeV, relative to using CFT information only to measure p_T.

3.3.2 Impact Parameter

Figure 24 shows the resolution of the impact parameter measurement from the L2STT using single muons. The dependence of the resolution on p_T can be parametrized by

$$\sigma_b(p_T)^2 = \sigma_0^2 + \left(\frac{a}{p_T}\right)^2 \tag{4}$$

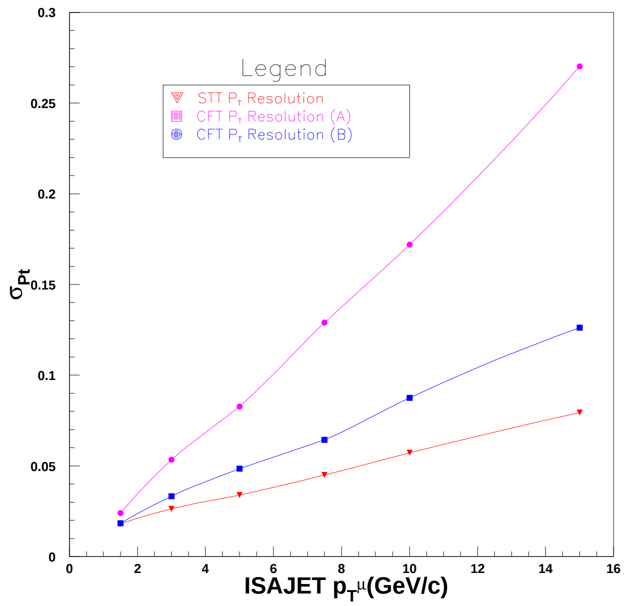

Figure 22: p_T resolution for single muons in the level-1 and level-2 trigger.

with $\sigma_0 = 18\ \mu$m, and $a = 53$ GeV$\cdot\mu$m. This parameterization also reproduces the impact parameter resolution for tracks from the primary vertex in $b\bar{b}$ and $t\bar{t}$ events.

We define the impact parameter significance

$$S_b = \frac{b}{\sigma_b(p_T)} \tag{5}$$

using $\sigma_b(p_T)$ from equation 4. As shown in Fig. 25, the distribution of impact parameter significance for tracks coming from the primary vertex is satisfactorily described by a Gaussian for both $b\bar{b}$ and $t\bar{t}$ events, although there are non-Gaussian tails. The width is very close to 1 for $b\bar{b}$ events, with a slight degradation and more pronounced tails in $t\bar{t}$ events, due to the higher track multiplicity. Furthermore, in both $b\bar{b}$ and $t\bar{t}$ events, the impact parameter significance distribution for tracks from the primary vertex is significantly narrower than that for all tracks (see Fig.26).

3.4 Beam Alignment and Stability

3.4.1 Initial Alignment of SMT with Beam

In the L2STT, 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.

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

$$\sigma(z) = \sigma(0) \cdot \sqrt{1 + \left(\frac{z}{\beta^*}\right)^2}, \tag{6}$$

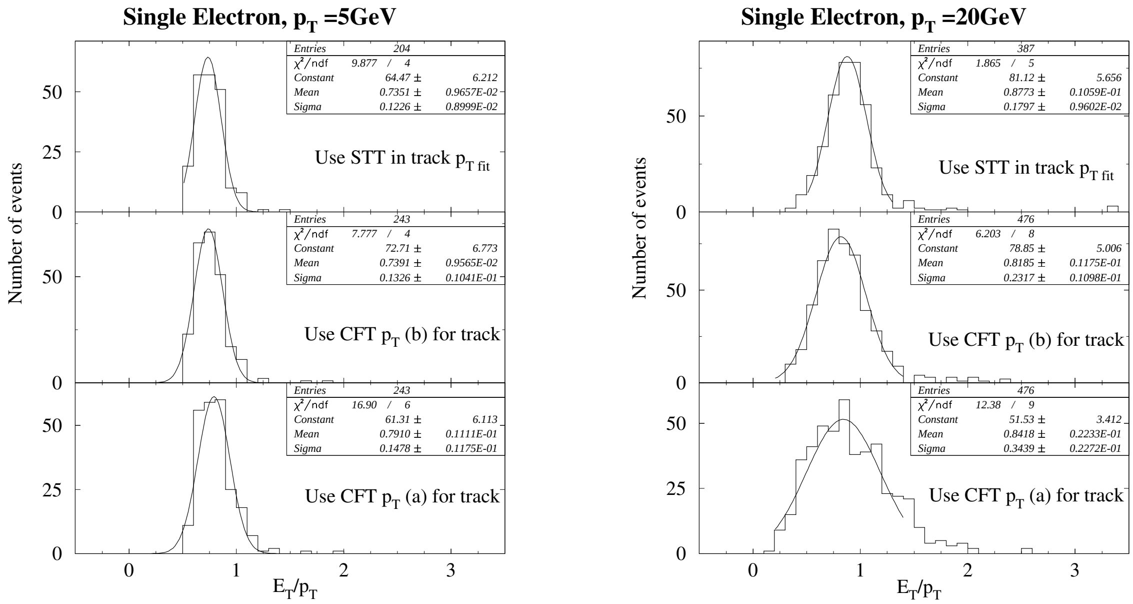

Figure 23: E_T/p_T for electrons with $p_T = 5$ GeV (left) and $p_T = 20$ GeV (right), assuming a calorimeter energy resolution of $13\%/E_T^{1/2}$.

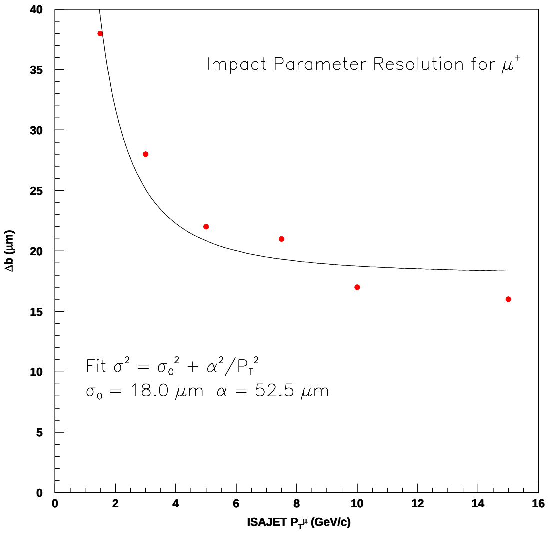

Figure 24: Impact parameter resolution vs p_T for single muons.

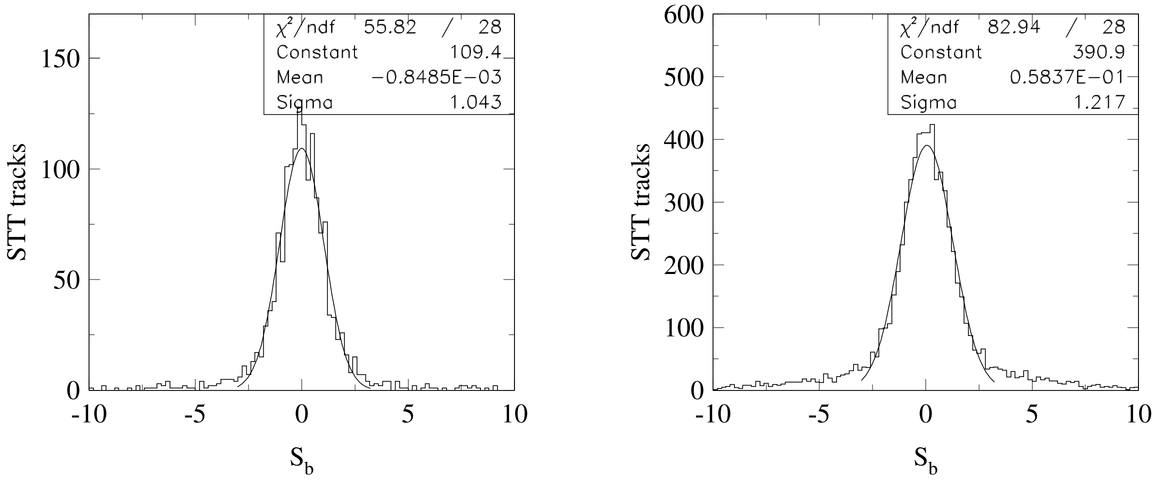

Figure 25: Impact parameter significance distribution for tracks matched to MC tracks coming from primary vertex, for $b\bar{b}$ events (left), and for $t\bar{t}$ events (right)

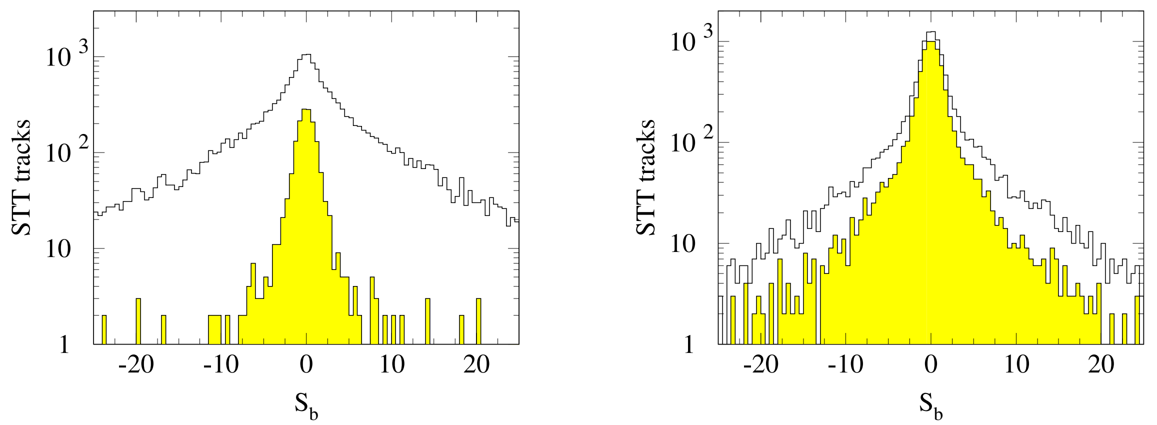

Figure 26: Impact parameter significance distribution for matched tracks, for $B \to K_s + J/\psi$ events (left) and $t\bar{t}$ →all-jets events (right). The shaded histograms correspond to tracks from the primary vertex.

where $\sigma(0)$ is the beam size at the nominal interaction point, and z is the distance from the nominal interaction point. During Run II $\beta^* = 35$ cm and $\sigma(0) = 40.5$ μm are expected. The longitudinal profile of the beam is Gaussian with a width of 35 cm, resulting in a Gaussian distribution of primary vertices along the beam with an rms of 25 cm.

Using these beam profiles we can compute the effect of a tilt (angular misalignment) of the beam with respect to its nominal trajectory. Figures 27–29 show the effect on the impact parameter distribution. Table 8 summarizes the effect of beam tilt. For optimal operation of the trigger, the beam tilt angle should be less than 200 μrad, else the fraction of tracks with large fake impact parameters becomes large, and gives rise to unwanted triggers.

Any residual offset between the detector and beam axes should be less than 1 mm. A larger offset would cause inefficiencies from tracks that cross SMT sectors. To achieve this precision, an iterative alignment procedure is planned.

Tilt Angle	Mean(b)	Fraction with		
		$b > 100$ μm	$b > 200$ μm	$b > 400$ μm
0 μrad	32 μm	1%	0%	0%
100 μrad	35 μm	2%	0%	0%
200 μrad	43 μm	5%	0%	0%
400 μrad	66 μm	20%	2%	0%
600 μrad	92 μm	38%	8%	0%
800 μrad	118 μm	51%	17%	0%
1000 μrad	146 μm	58%	27%	3%
2000 μrad	218 μm	77%	57%	26%

Table 8: Fake impact parameter due to misalignments.

Before the detector is rolled into the collision hall, a nominal beam position is determined based on the position of the low-beta quadrupole magnets. This actual beam position is expected to be within about 1 mm of this nominal value. The detector is centered on the nominal beam position to a precision of about 0.5 mm. After the detector is in place the SMT can be used to obtain a precise measurement of the beam trajectory relative to the detector. Limited corrections (± 1 mm and ± 250 μrad) can be made to the beam trajectory through orbit tuning. If the misalignment exceeds these margins either the detector or the low-beta quadrupole magnets have to be repositioned.

3.4.2 Beam Stability

Any variations in the beam position must be kept below 30 μm (approximately the width of the beam) else they cause an increase in the trigger rate due to the apparent impact parameter. Variations in the relative angle of beam and detector axes must be kept below 200 μrad.

During Run I, CDF monitored the beam position using their silicon vertex detector. The average offset of the beam axis from the detector axis was about 0.5 mm, and the average angle between beam and detector axis was about 500 μrad. 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 about 100 μm horizontally and about 150 μm vertically. Angle variations from store to store were typically less than 100 μrad, and of the order of 200 μrad over the course of the run. Variations during a store were much less important, typically 40 μm horizontally and vertically. Angle variations during a store were too small to be measurable [26].

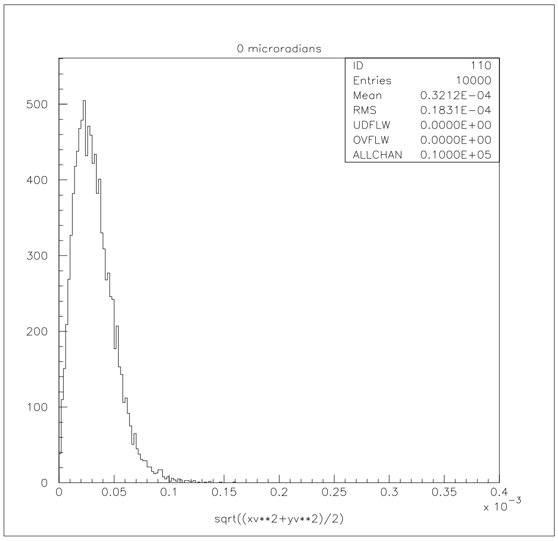

Figure 27: Fake impact parameter distribution without tilt angle.

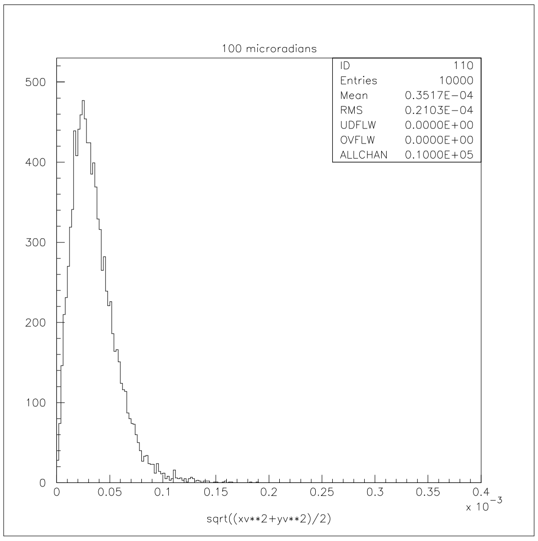

Figure 28: Fake impact parameter distribution for tilt angle=100 μrad.

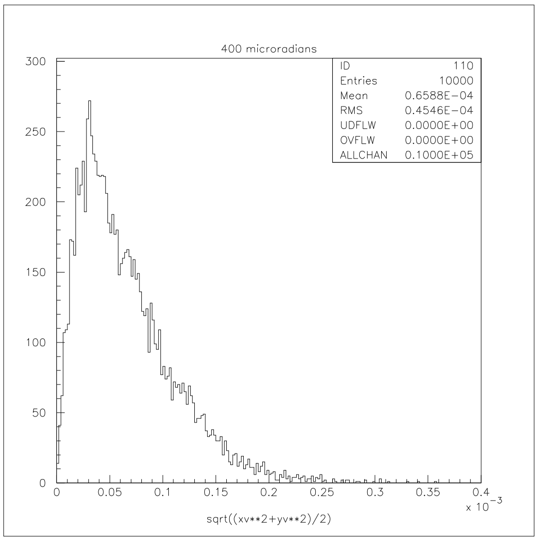

Figure 29: Fake impact parameter distribution for tilt angle=400 μrad.

In order to compensate for variations in beam position, CDF in collaboration with the accelerator division, has implemented a feedback system, using the measured beam position to set dipole corrector magnets near B0 to stabilize beam position and angle [27]. 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 μrad. 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 about 5 μm, and the beam angle to about 50 μrad. In a further test in February 1996 the beam was moved 50 μm vertically and 50 μm 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 the beam can be controlled independently at D0 and B0.

Similar correction magnets exist at the D0 interaction region and DØ will implement a similar feedback system to maintain stability of the beam position to the required precision.

3.5 Primary Vertex Determination

In level-3 the p_T of jets and electrons is computed from their energy E and shower centroid position z, measured using the precision readout of the calorimeter, and the position of the event vertex z_v as

$$p_T = E\frac{r}{\sqrt{(z-z_v)^2+r^2}}, \tag{7}$$

where r is the distance of the calorimeter from the vertex in the direction transverse to the beam. The luminous region of the Tevatron is expected to have an rms width of about 25 cm [28]. If no measurement of the z-position of the vertex is available, the uncertainty of 25 cm in the value of z_v would dominate the resolution for the p_T of jets and electrons at level-3 and wash out the trigger threshold. At level-2, a measurement of z_v is not required, because the p_T resolution is dominated by the fast trigger pickoff of the calorimeter.

We study these effects using Monte Carlo simulations [22]. The resolution of the z_v measurement as described in section 2.7 is expected to be 550 μm for events with one high-p_T interaction. For events with two additional interactions, the resolution is 610 μm. For a sample of $Z \to b\bar{b}$ events without additional interactions and with the event vertex within $-25 < z_v < 25$ cm, the efficiency for reconstructing a vertex is 98%. If we superpose two additional interactions on these $Z \to b\bar{b}$ events, the efficiency of reconstructing at least one vertex is 99%. The efficiency for identifying the correct vertex as the high-p_T vertex is 78%.

Due to the limited acceptance of the SMT (only the four inner barrel sections are equipped with 90° strips), the reconstruction of primary vertices with this method is limited to ± 30 cm around the center of the detector. About 30% of the high-p_T events have their interaction vertex outside this region. For these either no vertex is found or the vertex from a soft $p\bar{p}$ interaction inside this acceptance region is misidentified as the high-p_T vertex.

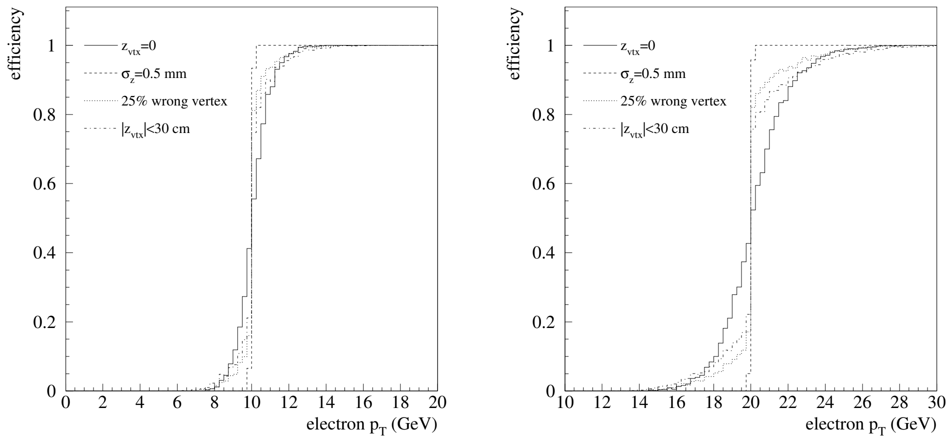

Figure 30: Turn-on curves for electrons with $p_T = 10$ GeV (left plot) and 20 GeV (right plot). For the solid curve $\delta(z_v) = 25$ cm is assumed, for the dashed curve the vertex is reconstructed with 100% efficiency, for the dotted curve the vertex reconstruction efficiency (78%) and the limited detector acceptance are also taken into account.

The dependence of the p_T trigger turn-on on the knowledge of the vertex position is shown in Fig. 30. Even if the correct vertex position is only found for $|z_v| < 30$cm (dash-dotted curve in the figure), the p_T threshold still sharpens significantly. Rates decrease exponentially with increasing threshold. Hence, a sharper threshold allows for a higher effective threshold value and a lower background trigger rate. This improvement implies that we can trigger more efficiently and with lower rates on single electrons at level-3.

4 Simulation of Physical Processes

4.1 Trigger Simulation

To simulate the performance of the L2STT, we generate events from processes of interest using the event generators ISAJET[29], PYTHIA[30], or COMPHEP[31]) and simulate the detector response using the GEANT program. The detector model also includes a detailed model for the charge transport mechanism in the silicon [32] and has been tuned to match test beam results. For all simulations we assume that the interaction vertex position along the beam direction (z_v) was normally distributed with mean(z_v)=0 and rms(z_v)=25 cm, as expected for Run II. We find that the expected transverse size of the beam of 30–40 μm has only a small effect on the results[33] and therefore assume infinitely narrow beams. All simulations assume a luminosity of $\times 10^{32}$ cm^{-2}s^{-1} and 132 ns between beam crossings.

We estimate signal efficiencies and trigger rates with and without the L2STT using Monte Carlo simulations. Without L2STT, triggers at level-2 are based on information from the calorimeter, muon system, CFT, and the preshower detectors. The additional requirements that can be imposed by the L2STT are:

- $\geq N_p$ tracks with $p_T > T_p$;
- $\geq N_b$ tracks with $b > T_b$;
- $\geq N_s$ tracks with $|S_b| > T_s$,

where N and T are programmable parameters. The track reconstruction algorithm for the L2STT at level-2 used in these simulations corresponds to algorithm (1) in section 3.2 with the following exception: if $\chi^2/dof > 4$ (dof=number of degrees of freedom), the SMT hit with the largest contribution to χ^2 was dropped and the remaining hits were fit again. All tracks with hits in at least three SMT layers and $\chi^2/dof \leq 4$ were considered "good" tracks.

The background event sample used to determine the trigger rates consists of dijet events, generated in six p_T bins. For each p_T bin, samples were created with 1, 3, 5, or 7 additional interactions. The additional interactions were simulated by low p_T dijet events. These 24 samples are combined weighted by their respective cross section times the probability of seeing the respective number of interactions at 2×10^{32} cm^{-2}s^{-1} and 132 ns between beam crossings. At this luminosity there are on average 1.3 additional interactions in each event. The results for +0, +2, +4, +6 and +8 interactions, for which no events were generated, were found by interpolating the results from the generated files.

The signal efficiencies determined from these Monte Carlo simulations should be quite reliable. On the other hand, past experience showed that estimates of trigger rates based on Monte Carlo simulations tend to underestimate the rates because it is very difficult to simulate noise, accelerator backgrounds, and pile-up from additional soft $p\overline{p}$ interactions correctly. We therefore see the rates quoted here as guidelines but should provide enough contingency in the trigger bandwidth to absorb higher rates.

4.2 Top Quark Pair Production

While triggering on $t\bar{t}$ $\rightarrow$dilepton or e+jets decays is straight forward, the μ+jets and all-jets channels are harder to trigger on. It is important to measure these decays of the top quark in order to verify that there are no unobserved decay channels. To study the benefits of the L2STT for the selection of $t\bar{t}$ events [33], samples of 2126 $t\bar{t}$ $\rightarrow$all-jets events and 1175 $t\bar{t} \rightarrow \mu$+jets events were generated using ISAJET.

To trigger efficiently on $t\bar{t}$ →all-jets and $t\bar{t} \to \mu$+jets decays we have to require the logical "or" ($\cup$) of a series of trigger conditions:

1: ≥ 4 jets with $p_T > 15$ GeV;

2: ≥ 3 jets with $p_T > 10$ GeV and $\not{p}_T > 20$ GeV;

3: ≥ 1 jet with $p_T > 45$ GeV;

4: ≥ 2 jets with $p_T > 20$ GeV;

5: ≥ 1 muon with $p_T > 6$ GeV and ≥ 2 jets with $p_T > 10$ GeV and ≥ 1 jet with $p_T > 20$ GeV.

The change in signal efficiency and background rate for additional requirements based on the L2STT is demonstrated in Fig. 31 for trigger condition 5 and for the $t\bar{t}$ →all-jets sample. The results are similar for the $t\bar{t} \to \mu$+jets sample. About 8% of the events are lost when L2STT-based requirements are imposed due to the acceptance of the SMT along the z axis. The trigger rate can be reduced by a factor of 4 if ≥ 2 tracks with $|S_b| > 3$ are requested. The corresponding trigger efficiency is about 80%.

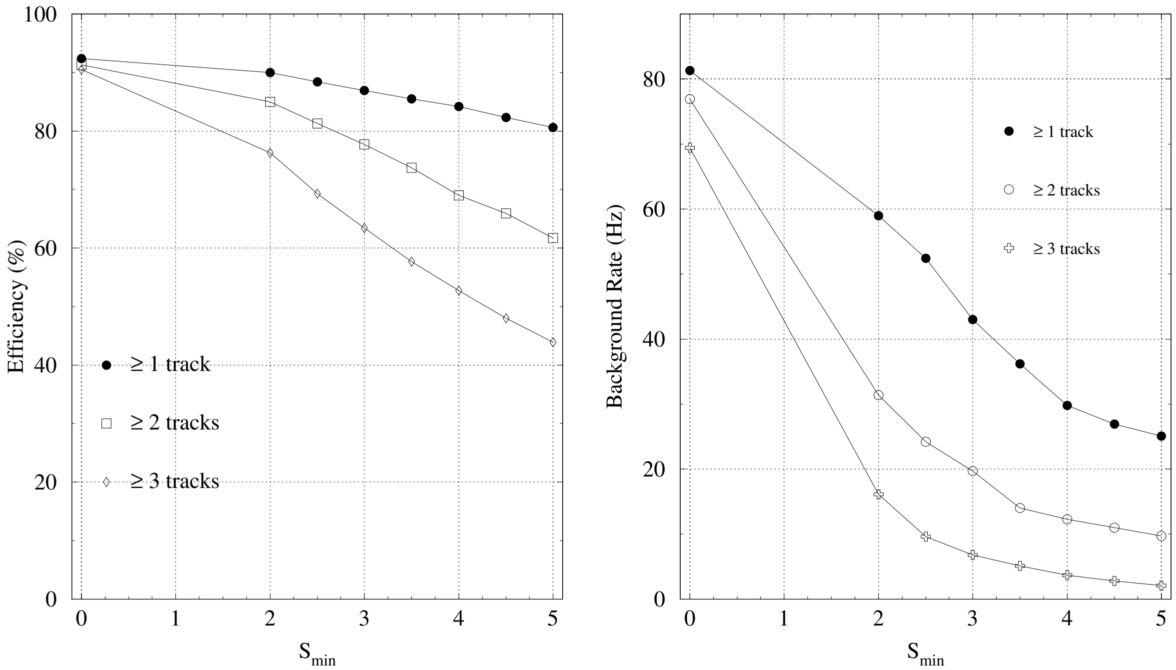

Figure 31: Signal efficiency for $t\bar{t}$ →all-jets events (left) and background rate (right) for the events satisfying trigger condition 5 together with the L2STT trigger condition $\geq N$ tracks with $|S_b| > S_{min}$.

We varied the transverse beam position to study the impact of finite beam size and found that smearing the transverse vertex position by 30 μm, corresponding to the beam width, leads to a 5% reduction in the signal efficiency for a fixed background rate. A systematic shift of the transverse vertex position in the of $\leq 100\ \mu$m does not introduce a signal efficiency loss as long as the L2STT criterion is adapted to maintain the background rate at a constant level.

Trigger conditions 4 and 5 account for most of the background rate but they are required to achieve high signal efficiency. If we require a logical "and" ($\cap$) of these two trigger conditions with the additional condition

6: ≥ 2 tracks in L2STT with $|S_b| > 2.5$

we can reduce the background rate substantially with minimal loss in signal efficiency. The rates and efficiencies for the two cases are listed in Table 9.

Table 9: Background rates and $t\bar{t}$ trigger efficiencies.

Trigger Condition	Background Rate	Efficiency	
		$t\bar{t} \to \mu$+jets	$t\bar{t} \to$all-jets
$1 \cup 2 \cup 3 \cup 4 \cup 5$	110 Hz	96%	100%
$1 \cup 2 \cup 3 \cup (4 \cap 6) \cup (5 \cap 6)$	40 Hz	93%	99%

4.3 Single Top Production

Single top production can lead to semileptonic finale states ($t \to \ell\nu b$) and hadronic final states ($t \to q\bar{q}b$). Again the hadronic decays are the most difficult to trigger on. The Monte Carlo event samples of for this process were generated with the COMPHEP generator. We select a combination of electron, muon and jet triggers with high efficiency for single top decays and investigate the effect of additional L2STT conditions on signal efficiency and background rejection [34]. The results for ≥ 1 track with $|S_b| > S_{min}$ are summarized in Table 10. For both leptonic and hadronic decay modes of the W, the L2STT allows a reduction of the trigger rate with little loss in signal efficiency. The L2STT rejections are substantial and necessary in view of the 1000 Hz bandwidth limit at level-2.

Table 10: Background rejection and signal efficiency for single top production.

Trigger Condition	Background Rate before L2STT	Additional Rejection		Additional Efficiency	
		$S_{min} = 2$	$S_{min} = 4$	$S_{min} = 2$	$S_{min} = 4$
1 electron with $p_T > 7$ GeV and ≥ 2 jets with $p_T > 7$ GeV	40 Hz	1.6	3.0	87%	74%
1 muon with $p_T > 6$ GeV	7 Hz	1.4	7.0	90%	78%
≥ 2 jets with $p_T > 15$ GeV	200 Hz	1.4	3.0	94%	80%

4.4 Associated Higgs Boson Production

We can easily trigger on associated Higgs production if the vector boson decays into electrons or muons. In order to have a chance at seeing a Higgs boson in Run II, we need to extend the acceptance of the experiment beyond the leptonic decays of the vector bosons and look at hadronic vector boson decays. If the vector boson decays into jets, triggering becomes much more diffcult.

To study this process, we generated a sample of 1700 $p\overline{p} \to WH^0 \to q\overline{q}b\overline{b}$ events using PYTHIA with $M_H = 90$ GeV [35]. We studied a number of calorimeter-based trigger conditions requiring between 3 and 5 jets with p_T thresholds between 5 and 15 GeV. We find that to maintain reasonable signal efficiency, we need to keep the required jet multiplicity and the p_T threshold low. A trigger requiring ≥ 3 jets with $p_T > 10$ GeV at level-2 achieves a signal efficiency of 74% at a background rate of 63 Hz.

Additional trigger conditions using track information offer the possibility to reduce the trigger rate while maintaining acceptable signal efficiency (see Table 11). For relative trigger efficiencies of 90%, the calorimeter trigger rate can be reduced by a factor of two to three. If we are willing to accept lower trigger efficencies, the effect of the L2STT becomes much more significant. At a relative efficiency of 80%, the L2STT triggers have rejection of a factor seven. From these results, it is clear that the L2STT can significantly help in triggering on $W/Z + H$ production. These conclusions also apply to techni-ρ decays ($\rho_T \to W\pi_T \to q\overline{q}b\overline{b}$).

Table 11: Background rejections for specific triggers and signal efficiencies ϵ for $WH \to q\overline{q}b\overline{b}$, relative to a trigger using only calorimeter information. The parameters N and X were varied to achieve the desired efficiency

Trigger Condition	$\epsilon \approx 90\%$	$\epsilon \approx 80\%$	$\epsilon \approx 50\%$
$\geq N$ tracks with $p_T > 3$ GeV	3.1	7.5	40
$\overline{b_{25}} > X$	3.0	6.4	26
$\overline{S_{25}} > X$	2.8	7.1	33

$\overline{b_{25}}$ - average impact parameter for tracks 2–5
$\overline{S_{25}}$ - average impact parameter significance for tracks 2–5
(tracks ordered in decreasing p_T)

4.5 Z Boson Decay to $b\overline{b}$

Z boson decays to $b\overline{b}$ serve as an essential control sample for the top quark mass measurement and the search for the Higgs boson. For this study [36], we use 10000 $Z \to b\overline{b}$ events generated with ISAJET. Most $Z \to b\overline{b}$ events have at least two jets with $p_T > 20$ GeV and we therefore define a set of calorimeter-based triggers which require the presence of at least two jets above a certain p_T threshold. For 40% efficiency the rate is about 260 Hz, 25% of the available bandwidth which is unacceptably high. Even for 20% efficiency the rate is 85 Hz, almost 10% of the bandwidth. To control the rate, we can use CFT information. For an acceptable efficiency of 35% out of level 2, the background rate still remains at about 20% of the bandwidth. Without additional rejection power this trigger can only be operated with a prescale and will not be able to exploit the full integrated luminosity.

Using the information provided by the L2STT provides the required additional background rejection. With the L2STT, we gain a factor of 2 in background rejection compared to CFT-based triggers for the same overall efficiency of 35%. For a background rate of as low as 20 Hz, we still have a signal efficiency of around 20%. This is five times higher than for the calorimeter-based trigger with the same background rate. Thus, the L2STT gives us the capability of triggering on $Z \to b\overline{b}$ decays at a low rate while maintaining an acceptable efficiency.

With a trigger efficiency of 20% and an estimated efficiency for tagging both b jets during the offline analysis of 25% we expect about 50000 signal events per fb^{-1}. Assuming the double

b-tag reduces the gluon-jet and light-quark background by an order of magnitude we expect 10^7 background events per fb^{-1}. For an integrated luminosity of 1 fb^{-1} the expected signal significance is thus about 17.

Table 12 summarizes these background rates and trigger efficiencies and Fig. 32 shows a plot of background rate versus efficiency for various trigger conditions relative to requiring $\geq$ 2 jets with $p_T > 15$ GeV.

Table 12: Background rates and signal efficiencies for $Z \to b\bar{b}$.

Trigger Condition	Background Rate	Efficiency
$\geq$ 2 jets with $p_T > 10$ GeV	1400 Hz	67%
$\geq$ 2 jets with $p_T > 15$ GeV	260 Hz	41%
$\geq$ 2 jets with $p_T > 20$ GeV	85 Hz	20%
$\geq$ 2 jets with $p_T > 15$ GeV and $\geq$ 2 CFT tracks with $p_T > 3$ GeV	200 Hz	35%
$\geq$ 2 jets with $p_T > 15$ GeV and $\geq$ 1 L2STT track with $\lvert S_b \rvert > S_{min}$	110 Hz	35%
$\geq$ 2 jets with $p_T > 15$ GeV and $\geq$ 1 L2STT track with $\lvert S_b \rvert > S_{min}$	20 Hz	20%

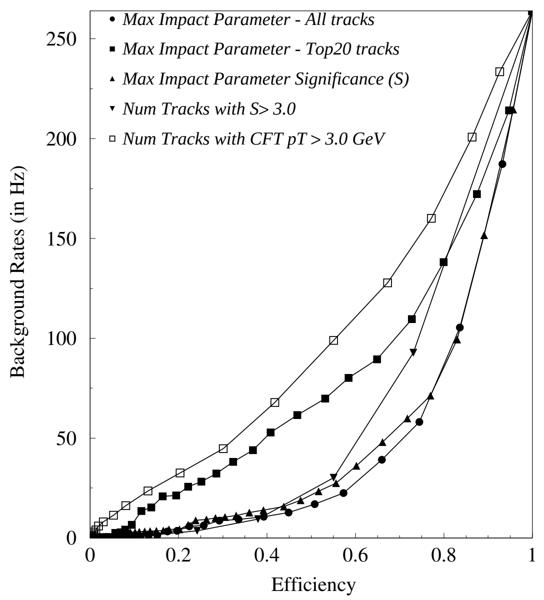

Figure 32: Background rate versus efficiency for $Z \to b\bar{b}$ decays relative to the trigger condition $\geq$ 2 jets with $p_T > 15$ GeV.

4.6 b-Physics

Although the b production cross section is large at the Tevatron, so is the background cross section. The challenge at a hadron collider is to isolate the b-physics signal from the copious backgrounds. Due to the event rate and bandwidth limitations in the data acquisition system, it is desirable to achieve as high as possible a signal-to-background ratio at the trigger level. For b-physics, this is best accomplished by using the L2STT to tag the long lifetime of b hadrons.

Since the L2STT is part of the level-2 trigger, any physics process must still trigger without SMT information at level-1 with a rate limited by the level-1 bandwidth ($\approx$10 kHz). Therefore, we do not consider the possibility of triggering on completely hadronic decay modes of b hadrons and concentrate on decay modes producing at least one muon in the final state.

A study of possible triggers for b-physics [24] considered three samples:

- $B_d^\circ \to K_s^\circ + J/\psi$, $J/\psi \to \mu^+\mu^-$ (CP asymmetry);
- $B_s^\circ \to D_s^- + 3\pi^\pm$, $D_s^- \to K^+K^-\pi^-$ (B_s mixing);
- $B_s^\circ \to D_s^- + \mu^+ + \nu_\mu$, $D_s^- \to K^+K^-\pi^-$ (B_s mixing).

In the B_s samples, the B on the other side was required to decay into $\mu + X$, resulting in at least one muon to trigger on. The background rate and signal efficiency for various trigger combinations are shown in Table 13.

For a trigger requiring a single muon with $p_T > 4$ GeV, the background rate is 40 Hz. We can reduce the rate by a factor 3 at a relative signal efficiency of 72% if we require ≥ 1 track with $|S_b| > 2.5$. By raising the threshold a similar rejection can only be achieved at the cost of a relative signal efficiency of 45%. The impact of the L2STT on the background rate and signal efficiency for $B \to K_s + J/\psi$ events is shown in Fig. 33. Using the single muon trigger with $p_T^\mu > 4$ GeV at level 1, the relative background rate and relative efficiency are shown for various L2STT criteria. A background reduction by a factor 10 can be achieved while preserving a 60% relative efficiency for the signal.

The second part of Table 13 shows the trigger performance obtained with various dimuon triggers. Asking for two muons with $p_T^\mu > 2$ GeV and at least one muon with $p_T^\mu > 4$ GeV provides an acceptable background rate but with a rather low signal efficiency (not shown in the table). Imposing a 2 GeV threshold on both muons leads to an increase of the 270 Hz background rate that is unacceptable at level-2. The L2STT would allow a significant reduction of this background rate while preserving good signal efficiency. This is illustrated on Fig. 33 where the trigger efficiency is shown versus the background rate for the condition ≥ 2 muons with $p_T > 2$ GeV+L2STT. Each closed circle corresponds to a different threshold on the track impact parameter. The L2STT allows reduction of the muon p_T threshold to 2 GeV in the dimuon trigger, so that it complements the single muon trigger. This is shown in the last line of Table 13: the logical "or" of the triggers 1 and 5 gives a total signal efficiency of $\approx$ 32% for the $K_s + J/\psi$ events with a level 2 background rate of 45 Hz.

The L2STT would also be very useful to control the rate of a low threshold dielectron trigger, allowing the collection of low-p_T $J/\psi \to ee$ events for calibration and physics analysis purposes.

Table 13: Background rate and signal efficiency for B decays for specific triggers combining muon and L2STT information.

	Trigger Condition	Background Rate	Efficiency $K_s + J/\psi$	$D_s^- + \mu^+$	$D_s^- + 3\pi^\pm$
single muon					
1	≥ 1 muon with $p_T > 4$ GeV	40 Hz	24%	12%	5%
2	≥ 1 muon with $p_T > 6$ GeV	14 Hz	11%	4%	1%
3	≥ 1 muon with $p_T > 4$ GeV and ≥ 1 track with $\lvert S_b \rvert > 2.5$	13 Hz	17%	7%	3%
dimuon					
4	≥ 2 muons with $p_T > 2$ GeV	270 Hz	20%	8%	0.5%
5	≥ 2 muons with $p_T > 2$ GeV and ≥ 1 track with $\lvert S_b \rvert > 2.5$	8 Hz	15%	6%	0.3%
6	$1 \cup 5$	45 Hz	32%	16%	4.9%

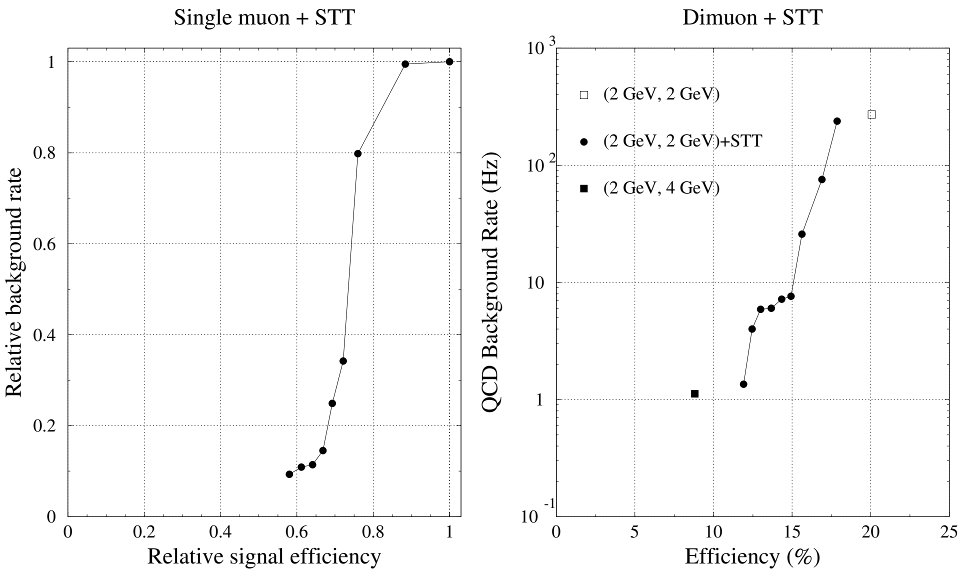

Figure 33: Background rate vs signal efficiency for $B \rightarrow K_s + J/\psi$ decays: using a single muon trigger with STT (left), and using a dimuon trigger, with and without STT, and with various p_t thresholds as indicated (right).

5 Cost Estimate and Schedule

We have performed a detailed cost estimate based on the conceptual design presented in section 2. We estimated the production cost of the fiber road card and the trigger card based on the cost of the VRB cards, which are of similar complexity plus the cost of the additional Altera 10k100 FPGAs, required for the trigger card. The cost of the track fit card is dominated by the funds required to purchase the processors. We assume 32 Texas Instruments TI320C62 digital signal processors (DSP) per crate, one DSP/track. For each of the three boards, we estimate prototyping costs equal to twice the production cost of the board plus $15k for test equipment. Table 14 shows the cost estimate for development and production of the full system. The engineering resources required for designing and testing the process were estimated based on previous experience with devices of similar complexity. The full engineering need is contained in the estimate, although we are seeking support for some of it from the universities and foreign institutions. One person-year (py) of an electrical engineer (EE) was estimated at $ 100k and one person-year of a technician (tech) at $ 70k. The total project cost (including 30% contingency) is about $ 1.9M.

Table 14: Cost estimate for L2STT project.

parts for	unit cost	number	spares	prototypes	total cost
fiber road card	$ 5200	6	2	2	$ 52k
trigger card	$ 7200	54	12	2	$ 490k
track fit card	$ 10000	12	2	2	$ 160k
vertex card	$ 5000	6	2	2	$ 50k
VME64 crates	$ 10000	6	0	0	$ 60k
Bit3 cards	$ 5000	6	0	0	$ 30k
test stands	$ 15000	3			$ 45k
total production cost					$ 887k
30% contingency					$ 266k
engineering for	electrical engineer		technicians		total cost
fiber road card	0.5 py		—		$ 50k
trigger card	2.0 py		—		$ 200k
track fit card	1.5 py		—		$ 150k
vertex card	1.0 py		—		$ 100k
testing	—		3.0 py		$ 210k
system integration	—		1.0 py		$ 70k
total manpower cost					$ 780k
total project cost					$ 1933k

The idea and design of this trigger processor has been extensively discussed within the DØ collboration [37, 1]. In April 1998, the collaboration conducted an internal review of the L2STT proposal. The review enthusiastically endorsed the proposal, concluding that this device was essential to the DØ collaboration's ability to carry out a fully implemented high-p_T physics program during Run II [38]. Since the review, the conceptual design has been further developed and we have secured engineering resources for the design of the device at the collaborating institutions. We are ready to begin the detailed engineering design as soon as we receive stage I approval for the project and funding for the engineers becomes available.

We can complete the first version of the design by spring next year and build first prototypes of all cards by early summer 1999. We expect to finalize the design, based on the experience gained with the prototypes, by October 1999. In March 2000, we are planning to conduct a system test with a small number of production boards to verify that the device works as a system. After that the remaining cards will be produced. By August 2000, the entire system will be at Fermilab and commissioning will start at DØ. The L2STT would therefore not be available at the beginning of Run II in April 2000. However, the accelerator will likely not begin running at the target luminosity at that time and the trigger system will be able to handle the rates without the L2STT. It is crucial that the L2STT become operational after the first shutdown, which is planned for August 2000. Table 15 summarizes this schedule.

Table 15: Schedule for development of L2STT.

date		milestone
April	1998	collaboration review of L2STT design
June	1999	prototype tests
October	1999	finalize design, begin production
March	2000	system test
April	2000	begin Run II
August	2000	installation at DØ
October	2000	L2STT operational

Four collaborating institutions will be responsible for the design and production of the L2STT hardware: Boston University (BU), Columbia University (CU), Florida State University (FSU), and State University of New York at Stony Brook (SUSB). We currently plan for BU to design and produce the trigger card. The Electronics Design Facility at BU already has experience with building trigger electronics for the muon system of the DØ upgrade. This design approaches completion and the engineers involved will direct their main efforts at the design of the trigger card for the L2STT. Students and a postdoctoral fellow from FSU will collaborate with the BU group in programming the FPGAs to cluster and filter the SMT hits. SUSB will have an electrical engineer and a postdoctoral fellow available for the design of the track fit card and the development of the track reconstruction algorithm. The CU electronics facility at Nevis Lab has significant experience in designing fast digital electronics and will be responsible for the design of the fiber road card and system integration. We are discussing with new members of the collaboration the possibility for their participation in elements of these, and in the vertex card.

These four institutions and the DØ collaboration have been actively seeking funding for the L2STT. In January 1998, the four institutions submitted a proposal to the Major Research Instrumentation program of NSF. It was not funded this year but will be resubmitted next year since reviews were favourable. In May 1998, a second proposal for part of the system was submitted to the NSF under the Knowledge and Distributed Intelligence program. During the past year, DØ has been succesful in seeking new international collaborators. The resources and funds that they contribute to the approved DØ upgrade, free up Fermilab funds that were committed to parts of the project that will be taken over by the new collaborators. These funds could be redirected to cover the cost of the L2STT.

The addition of the L2STT does not impact completion of the approved upgrade program. We specifically located the design centers for the device at universities, so that the design work on the L2STT does not draw Fermilab engineering resources away from the approved upgrade. The

L2STT does not require any changes in the design of the approved DØ upgrade. It fits seamlessly in the existing level-2 architecture, which allows additional preprocessor engines to be included. No component of the approved level-2 trigger system depends on the design of the L2STT, so that any delay that may occur in the design and production of the L2STT will not impact the readiness of the approved level-2 system.

6 Conclusion

The level-2 trigger system for the DØ experiment in Run II consists of an array of preprocessors which each process the data from one detector element. The only major detector system for which no preprocessor is foreseen in the approved DØ upgrade plan, is the silicon microstrip tracker. We have presented a design for such a preprocessor (L2STT) that fits naturally into the approved level-2 framework and enhances the power of the other level-2 elements. It does not require modification of any components of the approved upgrade.

We have demonstrated the benefits to the Run II physics program. The L2STT will make collection of enough $b\bar{b}$-events possible to be able to see a $Z \to b\bar{b}$ signal. This signal will be an essential control sample for top quark studies and for searches for new particles that decay into $b\bar{b}$, like the Higgs boson or technipions. The L2STT will reduce the trigger bandwidth needed to acquire all-hadronic final states that contain b quarks, such as $t\bar{t}$ →all-jets decays or Higgs boson decays associated with hadronically decaying W or Z bosons. It will sharpen the p_T threshold for track triggers at level-2 and for jet and electron triggers at level-3, which lead to a reduction of the rate of these triggers. The L2STT therefore reduces the level-2 trigger rates for a wide range of triggers. The bandwidth freed up in this way will benefit the entire Run II physics program.

A collaboration internal review has strongly endorsed the concept of the L2STT. The conceptual design of the L2STT presented has undergone significant scrutiny within the DØ collaboration. It is mature and ready for implementation. We have worked out a cost estimate for the project and we have secured engineering resources at collaborating universities to carry out the design. The collaboration is actively seeking funding for the design of the L2STT. Several proposals have been submitted to NSF. Funds contributed by new foreign collaborators may free up Fermilab funds for this purpose. At this time we are asking for stage I approval for the project, so that we can commit resources and start the engineering design.

References

[1] "A Silicon Track Trigger for the DØ Experiment in Run II — Technical Design Report", H. Evans, U. Heintz, T. Heuring, J. Hobbs, M. Johnson, S. Mani, M. Narain, F. Stichelbaut, H. Wahl, DØ note 3510 (1998).

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

[3] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. **80**, 3008 (1998) and FERMILAB-PUB-97/422-E, to be published in Phys. Rev.**D**.

[4] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **75**, 11 (1995) and Phys. Rev. **D52**, 4784 (1995).

[5] D. Karlen, presented at 29^{th} International Conference on High-Energy Physics, July 23–29, 1998, Vancouver, Canada.

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

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

[8] reference for CDF $Z \to b\bar{b}$.

[9] M. C. Smith and S. Willenbrock, Phys. Rev.**D 54**, 6696 (1996).

[10] T. Stelzer, Z. Sullivan, S. Willenbrock, Phys. Rev.**D 56**, 5919 (1997).

[11] "Standard Model Higgs at LEP", P. McNamara, invited talk at ICHEP98.

[12] The LEP Collaborations, the LEP Electroweak Working Group, and the SLD Heavy Flavour Group, CERN-PPE/97-154 (unpublished).

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

[14] B. Abbott *et al.*, submitted to Phys. Rev.D.

[15] "Technical Design Report for the Level-2 Global Processor", D. Edmunds, S. Gross, P. Laurens, J. Linnemann, R. Moore, DØ note 3402.

[16] `http://www-ese.fnal.gov/eseproj/svx/vrb/vrb.pdf`.

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

[18] "Dataflow Requirements for the Silicon Track Trigger", M. Narain, DØ note 3426.
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.

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

[20] "Cluster centroid algorithm for STT", U. Heintz, DØ note 3421.

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

[22] "Primary Vertex Finding al Level-2 using the Silicon Vertex Detector", M. Narain and S. Mani, DØ note 3338.

[23] "Timing studies of the STTpp", T. Heuring, DØ note 3492.

[24] "Triggers for B physics in run II", F. Stichelbaut, A. Zieminski, M. Narain, DØ note 3354.

[25] F. Carminati *et al.*, GEANT USERS GUIDE, CERN Program Library W5013, 1991 (unpublished).

[26] stability of beam at CDF in Run I ???.

[27] "Beam Motion and Feedback", P. Derwent, CDF note 3859.

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

[29] F. Paige and S. Protopopescu, BNL Report No. BNL38034, 1986 (unpublished).

[30] T. Sjöstrand, Comput. Phys. Commun. 82, 74 (1994).

[31] E. Boos, M. Dubinin, V. Ilyin, A. Pukhov, and V. Savrin, SNUT-94-116, hep-ph/9503280; A.P. Heinson, A.S. Belyaev, E.E. Boos, *Phys. Rev.* **D56** 3114.

[32] ref for SMT geant model ???.

[33] "Benefits of a silicon vertex trigger for $t\bar{t}$ events", F. Stichelbaut, DØ note 3312; "Update of the silicon track trigger studies for $t\bar{t}$ events", F. Stichelbaut, DØ note 3444.

[34] "Single top and the silicon tracker trigger", T. Heuring, DØ note 3400.

[35] "Towards a Run II trigger for $p\bar{p} \rightarrow WH^0 \rightarrow q\bar{q}b\bar{b}$", J. Hobbs, DØ Note 3321.

[36] "Triggering on $Z \rightarrow b\bar{b}$ events in Run II", M. Narain, DØ note 3339.

[37] "Working Group Report on a Vertex Trigger for Run II", F. Borcherding, M. Johnson, B. Klima, R. Lipton, S. Mishra, H. Piekarz, B. Winer, A. Zieminski, DØ note 2984; "Preliminary Hardware Design for a Silicon Vertex Trigger", M. Johnson, DØ note 3169.

[38] `http://www-d0.fnal.gov/cgi-bin/d0news?read_GENERAL_5865`