Readout Electronics Architecture and Packaging for the CDF Silicon Upgrade

Maurice Garcia-Sciveres *
Lawrence Berkeley National Laboratory

October 17, 1996

Abstract

The CDF Collaboration is replacing its silicon vertex detector (SVX') with an enhanced, completely new detector (SVXII). The SVXII design seeks to improve upon SVX' in many aspects simultaneously. Unlike SVX', SVXII information will be used for a second level trigger, demanding the ability to read out the detector at a very high rate. Additionally, the interaction period will now be significantly shorter than the lowest level trigger latency, which requires the use of a data pipeline in the front end. A desire for greater acceptance than SVX' confines the SVXII readout electronics to be directly on active sensor area. This demands small size (high interconnection density) and low mass, while at the same time the faster electronics call for improved thermal performance. New "deadtimeless" readout electronics and packaging to meet the above needs are discussed in this article. Preliminary results on performance of prototypes are presented.

1 Introduction

Run I of the Tevatron Collider at Fermilab ended on February 20, 1996, with a total integrated luminosity recorded by the Collider Detector at Fermilab (CDF) of roughly 110 pb^{-1} [1]. During this run, the vertex reconstruction

*Representing the CDF Collaboration

capability of CDF went from being in itself an experiment, to becoming the heavy flavor detection method that made possible the observation of the top quark and the measurement of its mass [2], as well as a wealth of b physics previously untapped in hadron colliders [3]. Vertexing in CDF began with the SVX detector [4], which was replaced after 19 pb^{-1} with a radiation hard "replica", SVX' [5]. For Run II of the Tevatron (to begin in 1999), CDF will replace SVX' with a new silicon detector, SVXII. Whereas SVX pioneered the use of a new technology in collider detectors, SVXII will seek to push this technology in several directions at once, so that while not as obvious, the challenge is equally great. SVXII will use 720 double-sided, 7.43cm long silicon sensors (width depends on layer), arranged in 5 coaxial cylinders (layers). A readout unit will consist of two sensors wire-bonded end-to-end with electronics on one of the sensors. Readout units will be mechanically paired, end-to-end with the electronics on the far ends, to form 29.8cm long "ladders". Ladders will be held in 5-layer "barrel" structures with 12 ladders per layer. Each barrel will have 2 beryllium bulkheads for mechanical support and electronics cooling. There will be 3 barrels end-to-end for a combined length of 87cm and a total of $405,504$ readout channels. The innovative features of SVXII presented in this article are (1) new "deadtimeless" electronics with fast digital readout that will permit implementation of a secondary vertex trigger, and (2) aggressive packaging to place the readout chips directly on active detector surfaces.

Table 1 compares some parameters of Run I relevant to silicon detector operation with the goals for Run II. The Tevatron luminosity will increase by more than an order of magnitude, primarily through the addition of more circulating bunches. Although possible, the luminosity per bunch will not be increased because the high $p\bar{p}$ cross section resulted in 2-3 interactions per beam crossing already towards the end of Run I. The large number of bunches means that there is no longer enough time between successive crossings for any kind of trigger decision. For Run II, therefore, all CDF detector elements must queue data from every beam crossing into a pipeline, waiting for the trigger processor to catch up. A 5.5 μsec trigger latency means that SVXII must remember 42 beam crossings. But perhaps the most ambitious aspect of the SVXII design is the desire to read out the detector at a projected rate of 40kHz in order to use SVXII information in a second level hardware trigger [6] The hope is to develop a hadronic b trigger based on vertexing [7]. A pair of ASICs jointly denominated svx3 will implement

Parameter	Run I	Run II
Beam crossing period	3.5 μsec	132 nsec
Trigger latency	3.5 μsec	5.5 μsec
Readout frequency	$<$10 Hz	40 kHz
Vertexing	2-D	3-D
Vertexing acceptance	60%	99%
Layer 0 radiation dose	0.3 MRad	1.5 MRad

Table 1: Comparison of Run I and Run II selected silicon detector parameters. The trigger latency refers to the lowest level trigger. Readout frequency refers to how often the silicon detector is fully read out. Layer 0 is the silicon layer closest to the beam pipe.

the necessary data pipeline and accomodate the high readout rate without compromising livetime. These chips are currently in the prototype stage.

In a hadron collider the interaction region is long in the beam direction ($\sigma_z \approx$ 30cm for CDF). This means that in order to achieve high acceptance a vertex detector must be long. However, the need for radiation hardness drives the active surface area of each readout channel to be as small as possible. Therefore, a strip detector must have short strips, making it necessary to place readout electronics within the active volume. In SVXII the electronics will be directly on the sensor surface, in order to avoid acceptance gaps. The packaging issue is being addressed in parallel with two different technologies: thick film on ceramic and multi-layer copper on kapton.

2 svx3 Chip Set

An svx3 chip set consists of two integrated circuits denominated Front End (svx3-FE) and Back End (svx3-BE), to be fabricated in a 0.8μm radiation hard CMOS process. The FE chip contains 128 channels, each with a charge integrator followed by an analog pipeline, while the BE contains 128 8-bit Wilkinson ADCs followed by a parallel-input serial-output memory with sparsification, and differential current drivers for an 8-bit output bus. The two-chip design attempts to isolate sensitive analog functions on the FE chip, in order to minimize induced noise from fast asynchronous switching required for digitization and readout. Both chips contain a shift register used to program several settings at run time, such as the length of the pipeline delay (d), an identification code for BE chip, and a sparsification threshold, among others. Unless otherwise noted, the rest of this section describes the circuitry of a single channel.

The FE charge integrator has a gain of 5mV/fC with a dynamic range of 450fC and an adjustable bandwidth for 10-90% rise times from 60 to 240ns with 30pF input capacitance. Full reset and settling takes 1.6μs, but the large dynamic range permits continuous operation over many beam crossings (resets will be carried out during gaps in the Tevatron beam structure, incurring no deadtime). Additional analog gain is supplied by the pipeline amplifiers (see below) for a total FE chip gain of 15mV/fC.

The FE pipeline is the circuit that permits both a long trigger latency and deadtimeless digitization and readout. It consists of a switched array of 47 capacitors, each of which can be placed in the feedback loop of one of two voltage amplifiers, the read (RA) and write (WA) amplifiers. A digital circuit shared by all channels, called the skip logic, determines which capacitor to connect to each amplifier given external strobe signals. The integrator connects to the WA input. 25ns Before every beam crossing, one of 46 "loop" capacitors is switched into the WA circuit, which is immediately reset to have zero charge on the capacitor for the present integrator output. After the beam crossing any change in the integrator is amplified onto the WA capacitor, which is then switched out and left to hold its charge. Capacitors are thus sequentially charged with correlated double samples of the integrator output. Upon a trigger, the capacitor from d crossings before the trigger is taken out of the loop by the skip logic. d is programmable between 0 and 41. Up to 4 capacitors can be out of the loop at any given time, and can remain out indefinitely. Capacitors out of the loop can be switched into the RA circuit whenever the BE chip is ready (see below), after which they are returned to the loop. The 47^{th} capacitor is not part of the loop, and is never in the WA circuit. It is placed in the RA circuit after every RA cycle as a reference in order to compensate for residual charges from switching. Due to the limit on capacitors that can be out of the loop, there is a maximum trigger rate achievable with 100% livetime, which depends on how fast the BE chip can return capacitors to the loop.

Figure 1 shows the integrator output (an internal FE signal) from a bench

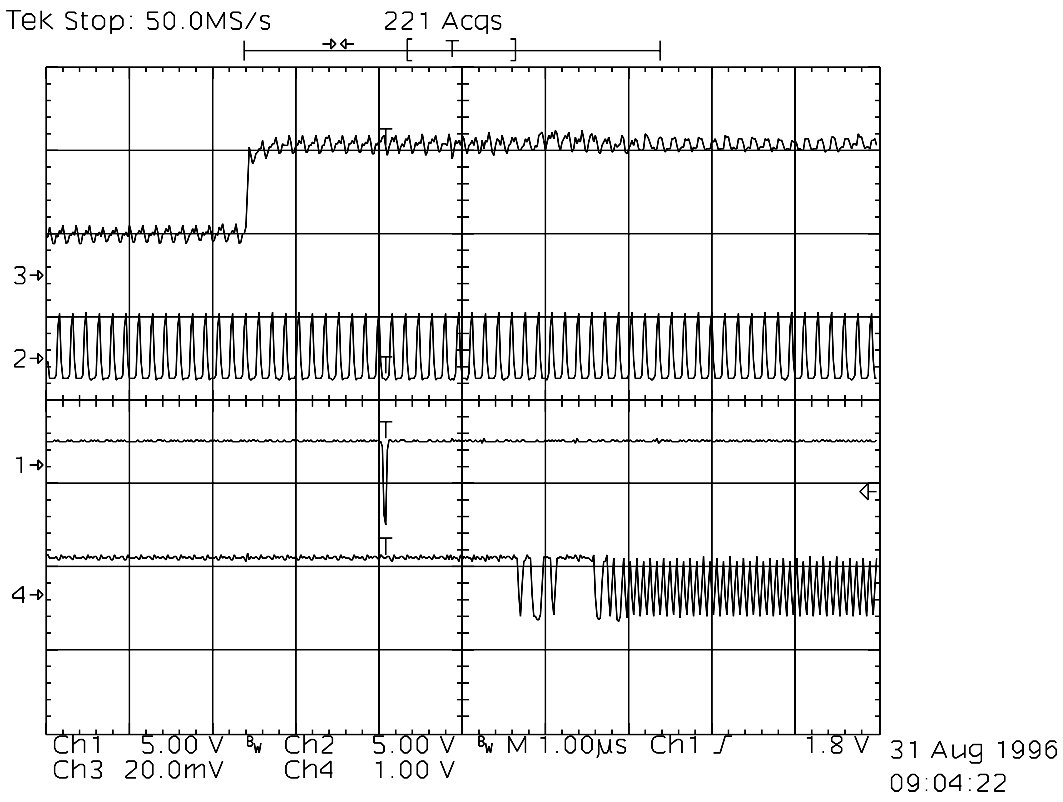

Figure 1: FE chip integrator output (trace 3) with beam crossing clock (trace 2), trigger (trace 1), and the BE clock signal (trace 4). Trace 3 was scaled by 0.1; the step corresponds to 200mV at the integrator.

test of svx3 prototype chips. A clock trace simulating beam crossings on the falling edge and a trigger trace can also be seen. The integrator ripples are partly an artifact of the setup for probing an internal signal, but there will always be some structure due to parasitic contributions from pipeline switching in the FE chip. However, all switching is internally synchronized with the beam crossing clock, so that it will be in phase with charge integration and contribute equally to all samples. On each rising clock edge a new capacitor is switched into the WA amplifier. A signal pulse at the integrator input (not shown) leads to the step at the output on the 15^{th} beam crossing, that will charge up the loop capacitor in the WA circuit at the time. 10 Clock cycles later a trigger occurs, that will take the charged capacitor out of the loop provided $d = 10$. d Should be programmed to match the trigger latency.

The FE chip circuitry ends with the RA output, which is to be digitized on the BE chip. A wire-bond connects the RA output to a BE chip ADC input. Digitization is based on a voltage to time conversion achieved by comparing the RA output to a ramping voltage [8]. This voltage ramp is generated in the BE and is common to all channels. The rate, direction and starting point are adjustable. Also shared by all channels is an 8-bit counter that advances on both edges of an external clock (BE clock). The counter measures time in integer counts, thus completing the analog to digital conversion. Channels are digitized in parallel: they all see the same voltage ramp and counter, but each has an independent comparator. When the comparator fires, the value of the counter at that instant is latched into a memory cell. After digitization the BE clock is used to sequentially strobe memory cells onto the output bus. Constant current differential drivers are used in order to reduce coupling of asynchronous signals to the FE integrator. Similarly, the BE clock uses a low level differential protocol.

The BE clock signal has been included in figure 1. Note that BE activity has no major impact on the FE integrator, as is imperative for successful deadtimeless operation. The chips under test were mounted on a prototype hybrid, but were only a single set (rather than several sets in tandem) and the hybrid was not mounted on and wire-bonded to a sensor. Thus figure 1 is not a conclusive demonstration of deadtimeless operation for SVXII, but it is an encouraging result.

Due to its parallel nature, digitization is quite fast. The design BE clock frequency is 53MHz for a 7-bit ADC time of just over 1μs. Prototype chips work with no problems at this frequency. Readout, on the other hand, is a sequential operation, as all chips on the same readout unit (half-ladder) share a single 8-bit bus. However, in order to use SVXII data for a trigger decision, the total readout time must be kept under about 3μs. With a readout rate of 53MBytes/s [9] this means that there is only time to read channels that register a signal (sparsification). Readout time is then determined by maximum occupancy, which is anticipated to be 10% at the inner layer. Worst case sparse readout time is predicted to be 2.3μs for the ladder with the highest occupancy (this includes reading the 2 nearest neighbors of every channel with a hit). Data are sparsified by a programmable threshold in the BE counter. When the counter exceeds the threshold, it produces a signal that flags all channels whose comparators have not yet fired. During sparse readout, only flagged channels (and optionally their neighbors) are strobed onto the output bus.

3 SVXII Readout Hybrids

Svx3 chip sets are mounted on sensors using hybrid packages. Hybrids provide power and signal distribution as well as heat dissipation. For deadtimeless operation there must be separate power planes and bypassing for the FE and BE chips. Thermal performance must be adequate to keep the sensors below 15°C. Additionally, because the electronics are in the detector active volume the hybrid mass must be minimized to prevent excessive multiple scattering.

Excluding the hybrids, the average material seen by particles traversing SVXII is estimated at 4.4% of a radiation length (RL). The data in figure 1 were taken using a prototype hybrid made of printed thick film on beryllium oxide ceramic (BeO). The thick film process used a fired tape dielectric that permits multiple layers with 100μm traces and spaces, as well as vias on a 250μm pitch. The process can combine gold traces with silver power and ground planes to reduce mass. Using BeO hybrids would add another 4.4%RL to the SVXII average material, which is deemed acceptable. Hybrids were also prototyped in a multi-layer copper on kapton flexible circuit (flex), which is laminated onto a beryllium substrate for heat dissipation. The process is limited to 4 metal layers and has a 500μm via pitch, which results in a larger surface area than BeO. Flex hybrids, including the Be substrates, would add 3.2%RL to the SVXII material. Part of the hybrid material is due to

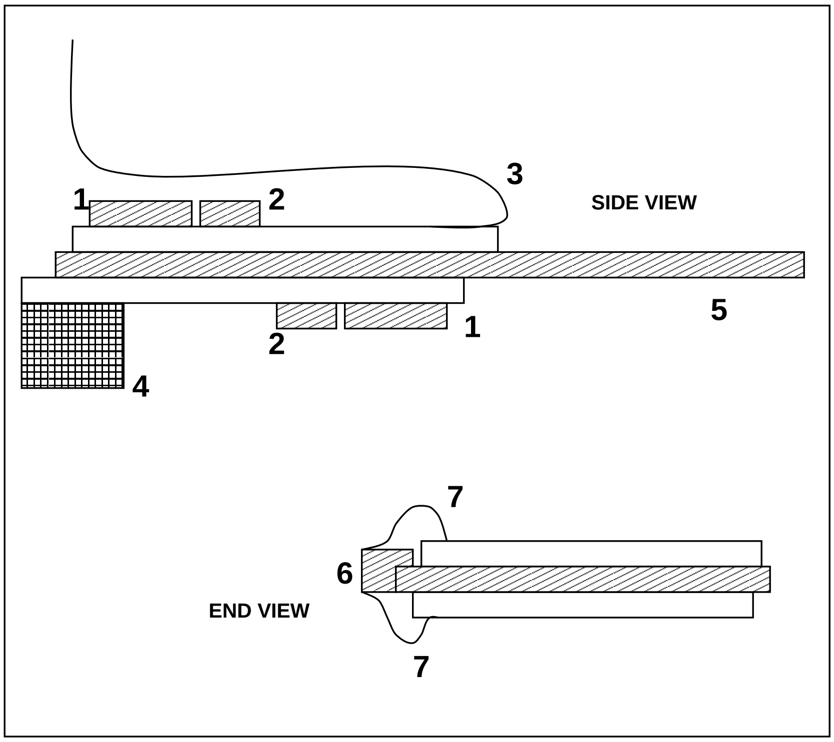

Figure 2: Schematic cross section of the electronics end of an SVXII readout unit or half-ladder. Only one of the two sensors in the half-ladder is represented (not to scale). Hatched shapes are silicon, open shapes are hybrids, and cross hatched is the bulkhead support/cooling channel. FE chips are marked **1**, BE chips **2**, readout cable **3**, bulkhead **4**, sensor **5**, jumper **6** (see text), and jumper wire-bonds **7**.

readout chips and other components, and is common to both technologies. Material is one of several hybrid properties that will be considered to choose a technology for SVXII. Other factors range from electrical performance to cost, schedule risk, ease of repair and ladder assembly.

Figure 2 shows schematic cross sections of the readout end of an SVXII ladder. The need to read out both sides of the sensor, as well as to support and cool the ladder forces the chips on the bottom (side view) far from the sensor end. This impacts on cooling, demanding a substrate thickness that increases rapidly with hybrid length. Given a power dissipation of 450mW per svx3 chip set (derived from measurements on prototypes) and a reasonably achievable coolant temperature of 0°C at the ladder, a flex (BeO) substrate thickness of 650μm (500μm) is required for a maximum sensor temperature of 15^oC at the ladder mid-point after a 1.5MRad dose [10].

Another challenge of placing readout hybrids on both sides of the sensor is making electrical connections to the bottom hybrid. A separate cable is mechanically not feasible without increasing layer radii, as well as electrically problematic because both hybrids read the same sensor, which favors tying them into a compact circuit. Top and bottom flex hybrids can be manufactured as a single unit and then folded over to cover both sides of the sensor, but this complicates ladder assembly. A conventional flex cable connecting the top and bottom hybrids is not possible because side-by-side ladders in the same layer must be very close to produce overlapping strips essential for off-line internal alignment (typical gaps are of order 2mm). A new solution being developed uses a very small (1.8mm $\times$ 1.5mm $\times$ 0.6mm) L-shaped jumper glued to the edge of the sensor (see end view of figure 2) that has 100μm diameter vias routing all connections from bonding pads on the top to pads on the bottom. Hybrids have a field of pads on one side for wire-bonding to the jumper. 10μm traces are used to scramble the connections on the jumper so that identical hybrids can be used on both sides of the sensor. The jumper is made out of silicon using micro-machining techniques [11], and has one layer of gold thin film on each side. Processing is expensive, but many parts are made simultaneously on a single silicon wafer, leading to low unit cost. These jumpers are currently being prototyped.

References

[1] The average live time was approximately 80%.

[2] F. Abe *et. al.*, Phys. Rev. Lett. **74** (1995) 2626.

[3] M. Paulini, 'HIGHLIGHTS OF B PHYSICS AT CDF', FERMILAB-CONF-**95-253-E** (1995); to appear in Nuovo Cimiento under Proceedings of 6th International Symposium on Heavy Flavor Physics, Pisa, Italy, Jun. 1996.

[4] D. Amiedei *et. al.*, Nucl. Instrum. Methods **A350** (1994) 73-130.

[5] S. Cihangir *et. al.*, Nucl. Instrum. Methods **A360** (1995) 137-140.

[6] D. Amiedei *et. al.*, Nucl. Instrum. Methods **A269** (1988) 51-62.

[7] presently, the only way to distinguish b events from low mass QCD on-line is to trigger on electron or muons from semileptonic b decays. Efficient off-line separation is not feasible because at $\sqrt{s} = 2$ TeV b events are relatively rare: $Br(p\bar{p} \rightarrow bX) = 10^{-3}$.

[8] The description of digitization has been simplified for clarity. Details of how voltages are presented to the comparator and of how the RA output is compensated using the 47^{th} pipeline capacitor have been left out.

[9] Two bytes are used to read out each channel, one for channel number and the other for ADC counts.

[10] Based on a finite element model that takes into account heat generated in the silicon sensors by a temperature-dependent leakage current. Thermal runaway does not occur in the model until the maximum sensor temperature approaches 25^oC.

[11] Micro-machining is a term commonly used to encompass a host of methods for chemical milling of silicon, making use of processing capabilities of the integrated circuit industry.