

Front End Busy Task Force Report

Fred Borcharding, Aurelio Juste, Eric Kajfasz, Don Lincoln, Ronald Lipton, Marvin Johnson, Yaroslav Kulik, Michael Weber

June 3, 2003

1 Introduction and sources of FEB

The D0 Run 2a DAQ system was initially designed for L1 accept rates of 10kHz. The experiment currently can run at 1.1 kHz with ~3-4% deadtime. The Front End Busy task force was charged with understanding SMT front end busy sources and recommending possible solutions to the shortfall in level 1 accept rate. This note is a summary of our studies, which were extended to examination of CFT front end busies and the front end busy components associated with L2 monitoring and services.

If we assume the maximum acceptable FEB fraction is 5%, the maximum L1 accept rate, r , is:

$$r = (0.05 - \text{fixed dead fraction}) / (\text{dead time per event})$$

where the fixed dead fraction is the rate-independent fraction of time the experiment is dead due to monitoring and other overhead. The rate calculation is particularly simple because the level 1 tracking electronics (CFT and SMT) are unbuffered. The SVX front end pipeline input stops while SVX data is digitized and read out to the VRB. In addition L1 accepts must be suppressed after SVX readout is complete until the SVX is in acquire mode for a period corresponding to the trigger latency (~9*396ns). We will refer to this as the "pipeline refill" time. If this is not done the trigger framework would take L1 accepts which correspond to crossings during the readout period while the SMT front end clock is not running and the resulting data would be incorrect.

2 SMT

We can separate the sources of deadtime in the tracking electronics into a fixed component, due to SVX chip setup, digitization and pipeline refill times, and a event dependent component, due to the readout time of channels above threshold and their neighbors. Figure 1 shows a logic analyzer trace of an SVX readout cycle in "read all" mode. Our current sequencer and VRBC firmware produces the following times:

Function	Time (seconds)
Chip overhead	7.9×10^{-8} /chip
Digitization	7.1×10^{-6}
Transition to readout	2.0×10^{-7}
pipeline fill	7.4×10^{-6}

Table 1. Measured SMT readout times

As discussed below these times have not been optimized.

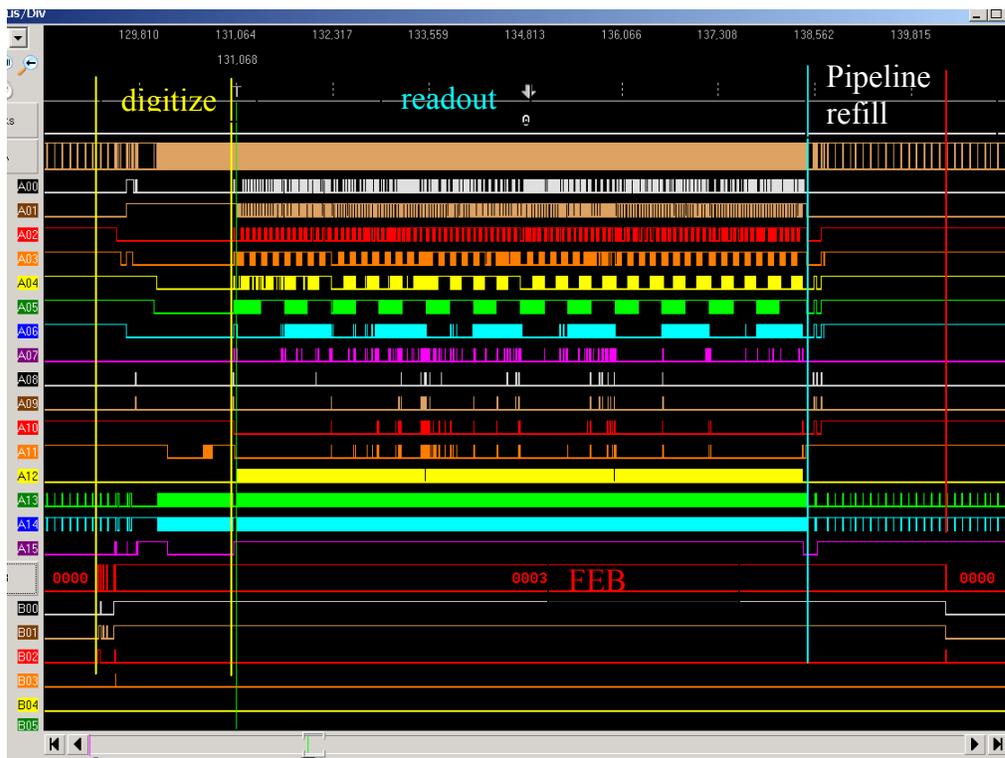


Figure 1. Logic analyzer trace of SMT readout showing digitization, readout and pipeline refill periods. The VRB-generated front end busy is the lower red trace. The SVX chip is in read all mode, with all 387 channels read out.

Figure 2 shows the distribution of readout FEB time as measured at the VRB controller output of an SMT (x65) and a CFT (x50) crate. This was measured in a recent run using an ORTEC time to amplitude converter and a QVT multichannel analyzer. The CFT crate is slightly busier than the SMT crate, as confirmed by DAQ monitor measurements. The SMT fixed dead time of about $14.6 \mu\text{s}$ is consistent with the times measured on the logic analyzer presented in table 1.

2.1 Digitization

SVX digitization consists of 1) A/D converter setup, 2) Digitization, and 3) pipeline collapse. The total time used from L1 accept to start of readout is $7.1 \mu\text{s}$. This includes $2 \mu\text{s}$ from L1 accept to start of digitize mode, $1.1 \mu\text{s}$ for ADC setup, and $1.6 \mu\text{s}$ for FIFO collapse. Actual 8-bit digitization takes $2.41 \mu\text{s}$. These times have not been optimized for speed. It is likely that with some effort they can be tuned to reduce the total digitization mode overhead *from 7.1 to 4.6 μs* . This will require careful tests that any changes in SVX sequencer timing do not affect data quality.

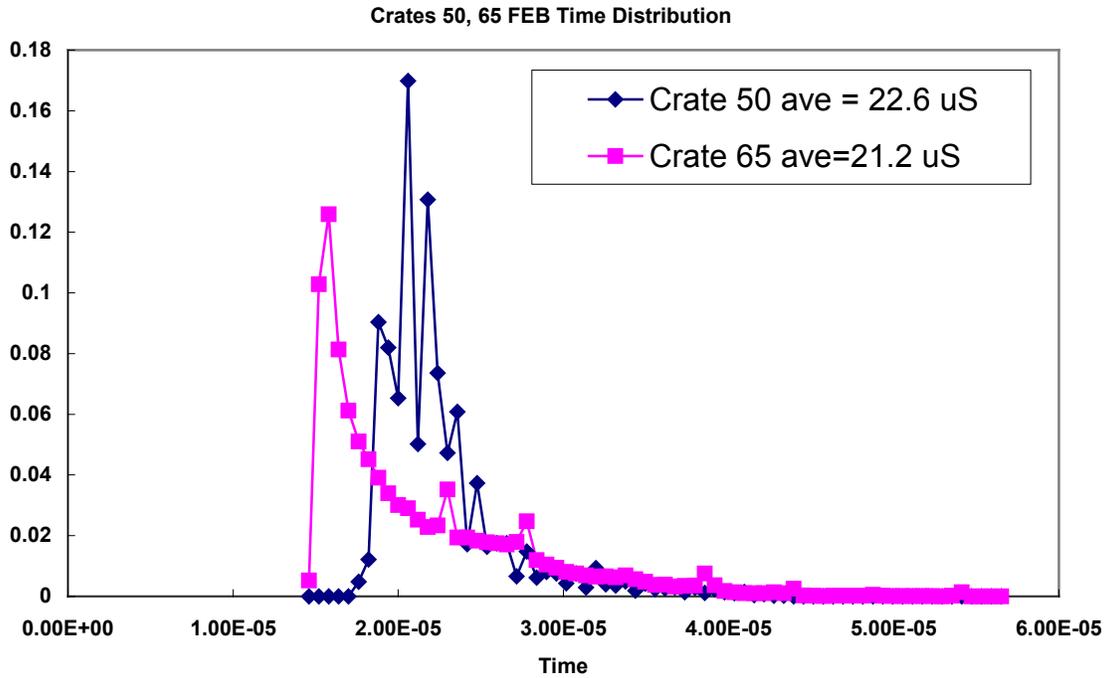


Figure 2. Front End Busy time distributions for a typical SMT and CFT crate.

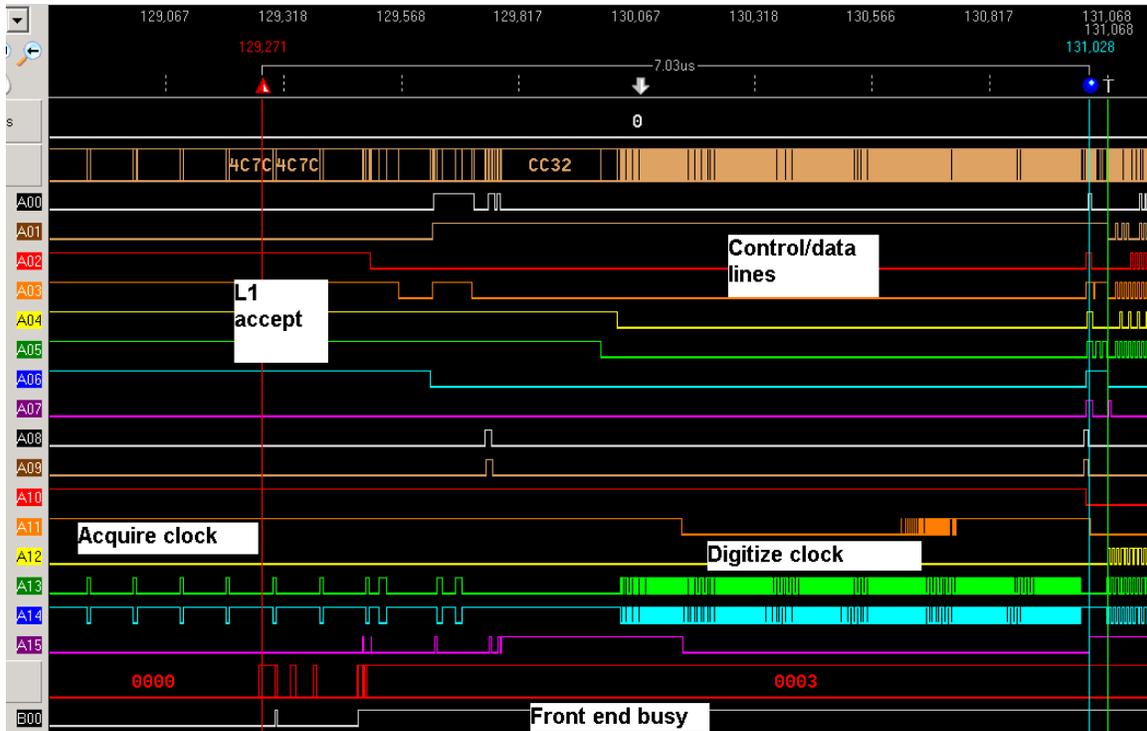


Figure 3. SVX digitization timing.

2.2 Pipeline refill

The trigger framework is designed to accept triggers whenever FEB is not asserted. After SVX readout the chip pipeline is reset and contains incorrect and out of time data. In order for the SVX to provide data that corresponds to the same crossing as the rest of the experiment, and to eliminate triggers caused by SMT readout noise, FEB must be held high for a full L1 latency period after acquisition mode is re-entered. This was not done properly for data taken before March 2003, but the effects were minimized by a few bad HDIs with long readout. This was corrected by adding a fixed additional FEB time in the VRB controller. This time is currently $\sim 7.4 \mu\text{s}$, much longer than the time needed to refill the pipeline. The time is currently also dependent on data load – due to the mismatch between the 50 MHz VRB clock and the 53 MHz readout clock. Proper setting of the pipeline refill time will require a careful understanding of cable delays and changes to VRB hardware or firmware to provide a fixed, data volume independent pipeline refill time from the VRB which corresponds to the trigger latency. *It is likely that this time can be decreased from 7.4 to $\sim 3.5\text{--}4.5 \mu\text{s}$.*

2.3 Data Load

Figure 2 shows that the current readout time distribution for a typical crate peaks at the minimum and falls rapidly. The average contribution of data to deadtime is $7.2 \mu\text{s}$ of the 21.2 total for the SMT. This picture is very different than the situation before March 2003, when the readout time was dominated by a few noisy HDIs and contributed $11.4 \mu\text{s}$ to the event readout time.

2.3.1 HDI monitoring and procedures

Much of the improvement in L1 rate capability can be attributed to improved monitoring of HDI performance by the SMT shifters. Stripcharts, such as the one shown in figure 7, indicate if one or more crates of the 12 SMT crates are generating large FEB. If this is the case it is usually easy to identify the HDI at fault by examining plots of the most occupied HDI in each event and the event-by-event occupancy monitor. The HDI can then be checked for pedestal position and stability and either be redownloaded or turned off until the problem is fixed.

2.3.2 Truncation

The SVX sequencer has the ability to truncate data readout at a fixed time. This truncation is intended to allow the SMT to terminate the readout of particularly noisy HDIs that can dominate the deadtime. This is based on the hypothesis that the long HDIs are dominated by coherent noise and very little real data is lost by truncation. Data truncation was simulated by cutting HDI output data beyond n hits. The distribution of the number of SMT hits per track as a function of the truncation is shown in figure 4. The hit/track efficiency begins to plateau around 200 hits/HDI where it is a 1.3% effect on the SMT hit count. We note that there is already a cut in the reconstruction where chips with more than 25% occupancy are not used in the tracking. A truncation value of 300 hits (0.3% effect on SMT hits) would decrease the number of hits/event from 420 to 280 and increase our current SMT level 1 rate capability by 300 Hz.

HDI cut off effect

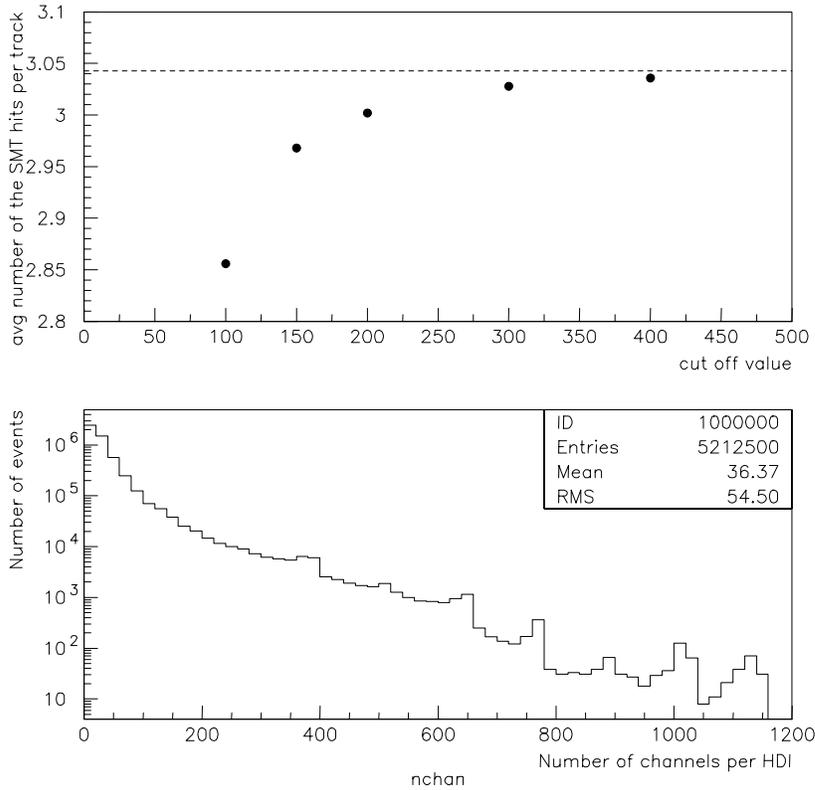


Figure 4. a) Number of SMT hits per track as a function of number of hits/HDI cutoff. b) distribution of number of hits per HDI.

The HDI hardware timeout is implemented in the SVX sequencer. The option can be turned on and off remotely but the time value must be set in the firmware. A version with a 25 μ s timeout was tested in April and caused frequent crate hangs at high L1 rates. This problem must be solved before the hardware timeout is used. It appears that the SVX data load is sufficiently under control with proper shifter monitoring and elimination of bad HDIs that implementation of truncation will not be necessary in the near future.

3 CFT / CPS / FPS

The systematic of the CFT readout are very similar to those of the SMT. We will concentrate on the items that are different in the two systems and how they can be improved. The CFT, CPS and FPS share the same hardware and can be treated similarly. Unless explicitly stated otherwise this section applies equally to all three.

As in the SMT system there is a fixed overhead that includes digitization, SVX Readout, and pipeline refill. There is also a variable length readout associated with the data load. Unlike the SMT system the AFE boards only use 64 of each 128 channels in each SVX chip. We have found that these non-used channels can turn up in the zero suppressed readout, which adds to the “fixed” data load. The length of the readout is dependent upon both the thresholds and the occupancy. Improving thresholds can shorten the data length. Two problems associated with threshold improvement are the phenomena of pedestal oscillations and double peaked pedestals.

Function	Time (seconds)
Digitization	7.84×10^{-6}
L1 to start Dig	2.83×10^{-6}
to end Dig	5.01×10^{-6}
VSVX	3.4×10^{-6}
pipeline fill	4.4×10^{-6}
Unused channels	1.8×10^{-6}

Table 2. Measured CFT readout times

3.1 Fixed Overhead

Figure 2 shows that the CFT currently has a larger fixed overhead than the SMT (~17 vs ~14 μ s). The components are itemized in table 2. There are two items that appear in the CFT list which are not present in table 1. Data from the SIFT discriminators are read out as a “virtual SVX” appended to the string of real SVX data in the AFE. This represents a fixed data load of 130 bytes with associated readout time

The time from L1_Accept at HDI to end of digitize or start of readout is 7840 ns. Of that time 2830 ns is from L1 to start digitize, and 5010 ns is from start digitize to end digitize. New SEQC code exists that removes 3300 ns from this time. (*Note: Sequencer modifications tested on May 21 have successfully reduced the fixed overhead to ~13.6 μ s.*) The VSVX Readout is a fixed time of 3700 ns. The vsvx data is not zero suppressed and consists of 128 bytes of data + 2 bytes header = 130 bytes. Which take $130 \times 25 \text{ ns} = 3250 \text{ ns}$ to readout plus some 450 ns of setup and other times.

There is also some 1080 ns from the end of readout to start of acquisition. It may also be possible to save time here. The time for the Pipeline Refill must also be counted and for; our depth is 33 crossings or 4343 ns.

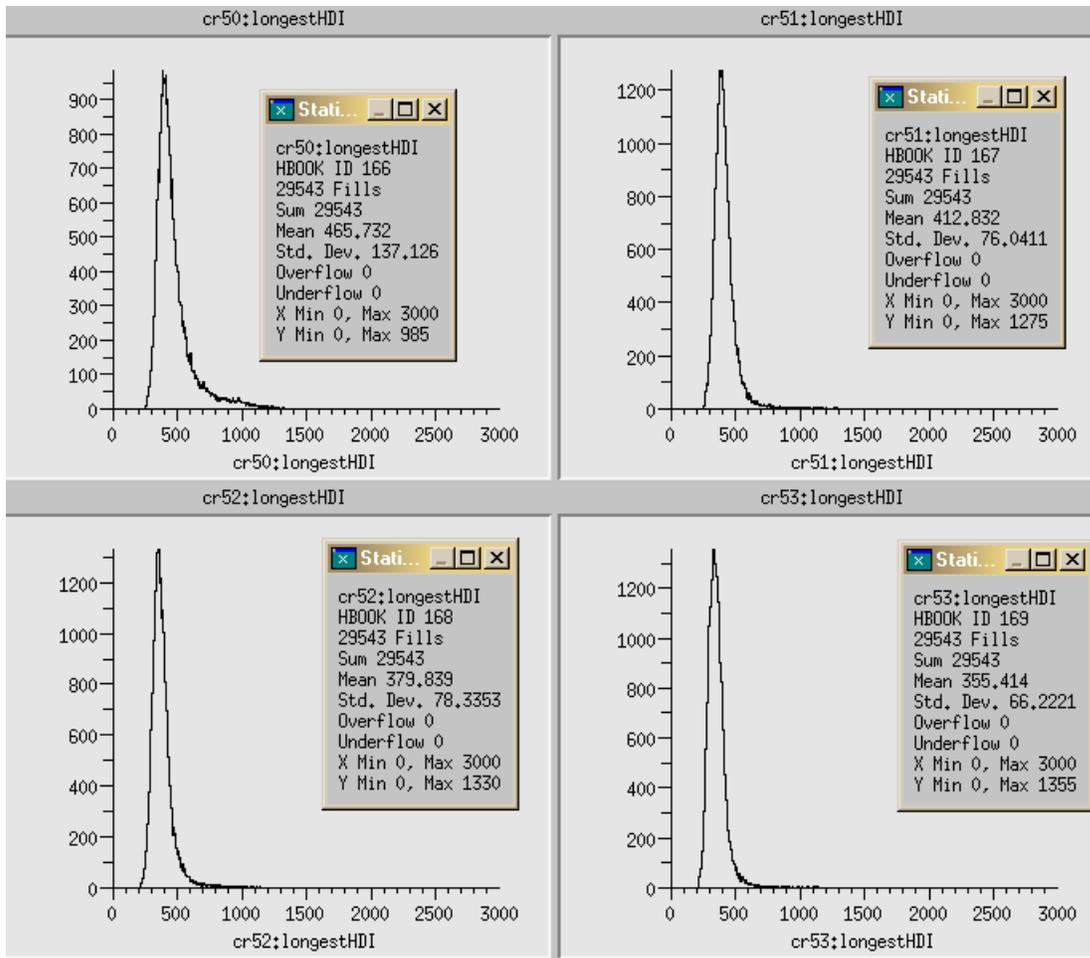


Figure 5. Histograms of the longest HDI for each event for each of the four GS crates. For each event the length of the longest HDI for the GS crate is entered into the histogram. The data is a Poisson distribution starting about 200 to 250 bytes and with means of from 350 to 470.

3.2 HDI length

The length in bytes of the longest HDI at a luminosity of 19E30 is shown in figure 5 for the four CFT crates. Crate 0x50 which contains the CFT Axial fibers is the longest of all crates. This contains data Readout from SEQ.

Fixed data lengths

- 130 bytes for vsvx
- 16 bytes for svx headers
- 2 bytes for SEQ header
- 32 bytes for VRB header VRB only
- 180 bytes total

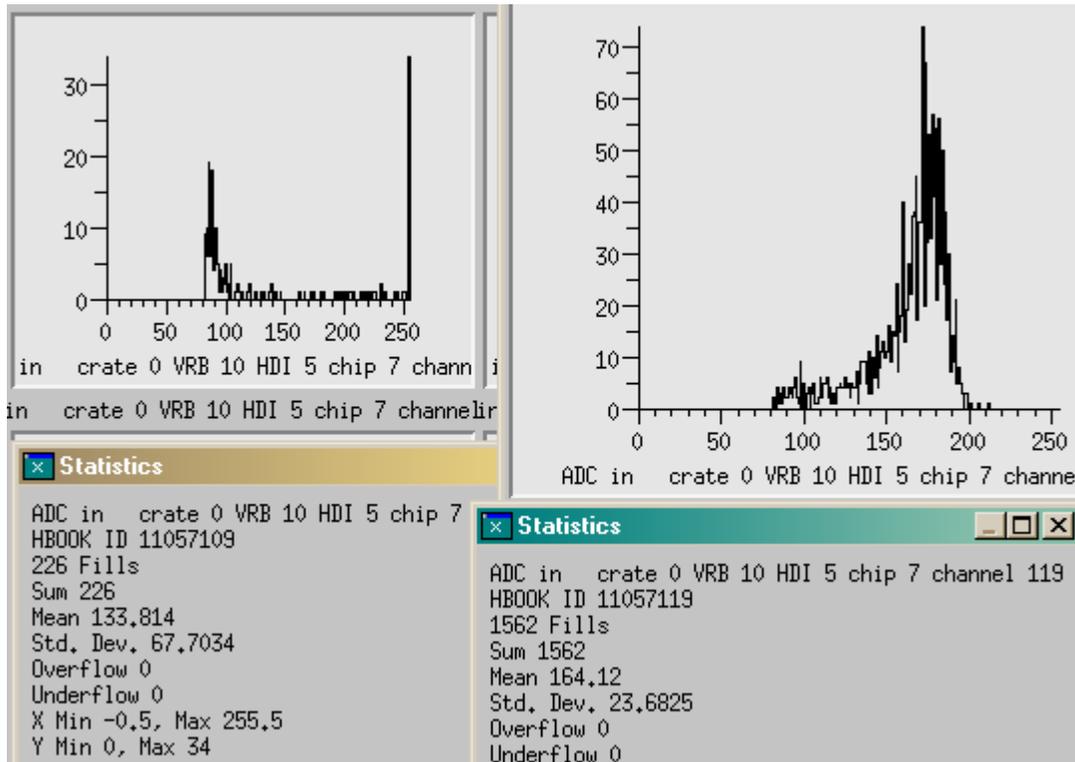


Figure 6. Histograms of the ADC spectra in counts for two CFT channels. On the left is a channel connected to an active fiber. On the right is a channel that is not connected to any fiber.

The longest HDI plots shown should start at 180 bytes, but instead begin at from 200 to 250 bytes for the different crates. These extra bytes are from the readout of the unused channels. The ADC spectra with colliding beam for channels that are bonded to SVX input from SIFT analog output are shown in figure 6. In a used channel a peak is seen just above the zero suppress value and the entries are dominated by overflows. In many of the unused channels the pedestal peak is well above the zero suppress value and thus that channel is present in most events. Because of the manufacture of the boards this problem is worse in those AFE boards with a green solder mask. These are the boards used for all of the CFT axial and most of the CFT stereo. This corresponds to all of crate 0x50 and most of crate 0x51.

The relative Occupancy of un-used channels can be estimated from the figure. The bonded are on left and the un-used are on right. The bonded plot has 226 entries, the un-used has 1562 entries, or about 7x the entries as bonded. So if we take

$$\begin{aligned}
 8 \text{ channels} \times 7 \text{ times more occupancy} &= 56 \\
 64 \text{ channels} \times 1 \text{ occupancy} &= 64
 \end{aligned}$$

We see that almost 1/2 of the total occupancy can come from the un-used channels. From figure5 the shortest data from for 0x50 is some 252 bytes.

$$252 - 180 \text{ fixed overhead} = 72 \text{ bytes are from unused}$$

This 72 bytes divided by 2 gives some 36 channels of unused per HDI, for an average of 4 per MCM out of a possible 8. This translates into

$72 \text{ bytes} \times 25 \text{ ns} = 1800 \text{ ns}$ of extra readout time.

Work is being done on a hardware fix to reduce to near zero the number of un-used channels in the readout data. This fix is presently in the testing phase. All the AFE boards on the platform will have to be removed and modified to completely implement this fix. A long access is the best way to make this fix, because it involves soldering 64 very small resistors on each board.

3.3 Reduction Options

There some options beyond those discussed above to further reduce the FEB. Each of these options involves some compromise of the physics data and can be phased in at different levels. These options include:

- move from 40MHz to 53MHz readout (installed by SEQ/VRB >> 8 AFE boards)
- Do not readout the VSVX at all (download to AFE boards can be changed from run to run)
- readout the VSVX only (only requires SEQC firmware change >> GS crate)

3.3.1 53MHz Readout

The HDI's are presently readout using a 40MHz clock. Going to the maximum frequency clock or 53MHz could reduce this readout time. This requires a hardware modification to each board that could be implemented in groups of one SEQ board or 8 AFE boards. This would provide an average savings of $\sim 1.4\mu\text{s}/\text{event}$ and is probably not worth the effort involved in hardware modifications. Note that this hardware modification can be done simultaneously with the unused channel modifications. Also it can be made backwards compatible to 40MHz readout.

3.3.2 Dropping the vsvx

Dropping the vsvx data from the readout has the least impact on the physics of these options. The present offline tracking algorithms do not use the discriminator data and depend only on the ADC values. The VSVX is used for CTT monitoring, debugging, and verification. We should consider dropping this readout only after the CTT has been fully debugged.

3.3.3 VSVX only

This option reduces the readout time because if the SVX is not read out the digitize time of over 4 us can be saved as well. Thus this option in principle results in the shortest readout time of any of the changes. The physics is maximally impacted by this option since the analog data from the CPS and FPS would no longer be available.

4 Level 2

All VRB based systems have 14 level 1 buffers which store L1 accepts until a level 2 accept/reject decision is made. When these buffers fill L1 busy is asserted until a level 2 decision provides additional free buffers. A level 2 latency which allows 14 L1 decisions

to stack up will cause coherent front end busies in all of the tracking crates independent of data volume. Figure 7 shows a “stripmon” chart of FEB fractions for the SMT and trigger framework crates. There is a clear spiked structure that appears coherently for all tracking crates, indicating filled tracking L1 buffers.

To determine the source of the spikes a set of runs were taken with various L2 conditions. Figure 7 shows the effect of turning level 2 monitoring off. Most of the spikes are gone and there is now an obvious baseline. Figure 8 shows the FEB fraction as a function of L1 accept rate for runs in store 2438 with monitoring off and on.

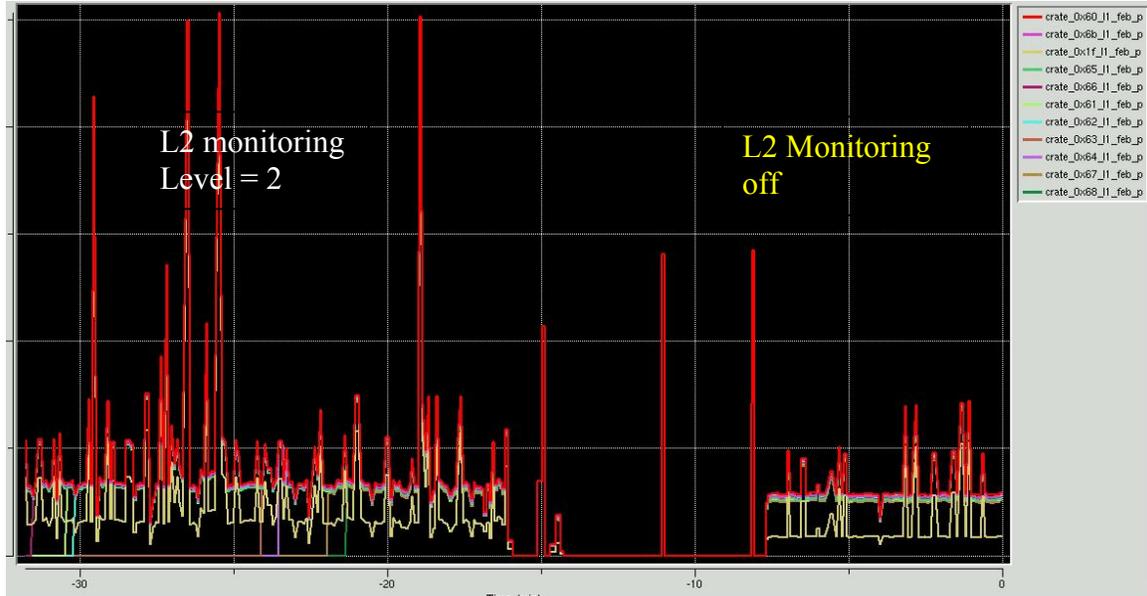


Figure 7. Front end busy fraction as a function of time for SMT and trigger framework (yellow) crates with level 2 monitoring on and off.

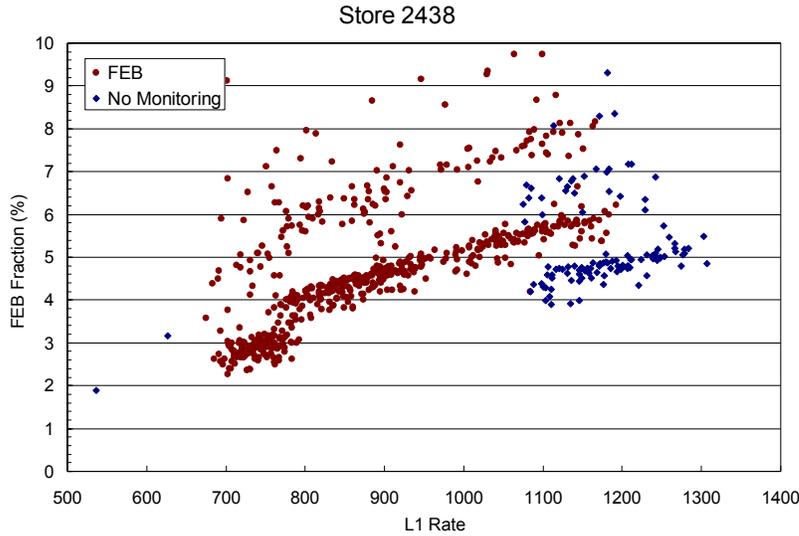


Figure 8. Overall front end busy fraction for store 2438 as a function of level 1 accept rate for data with monitoring on (brown) and off (blue).

There is about a 1.2% difference in FEB as well as a secondary peak due to the remaining spikes. Figure m shows the effect of turning both monitoring and level 2 services off. With all services (SSH access) off level 2 appears to have no contribution to L1 busy.

Slow VME transfers of histograms between the SLICs and the Alphas dominate level 2 monitoring time. The monitoring provides eight levels of detail. Until May of 2003 the experiment ran at level 2, which provides 7 output histograms. After a brief period of

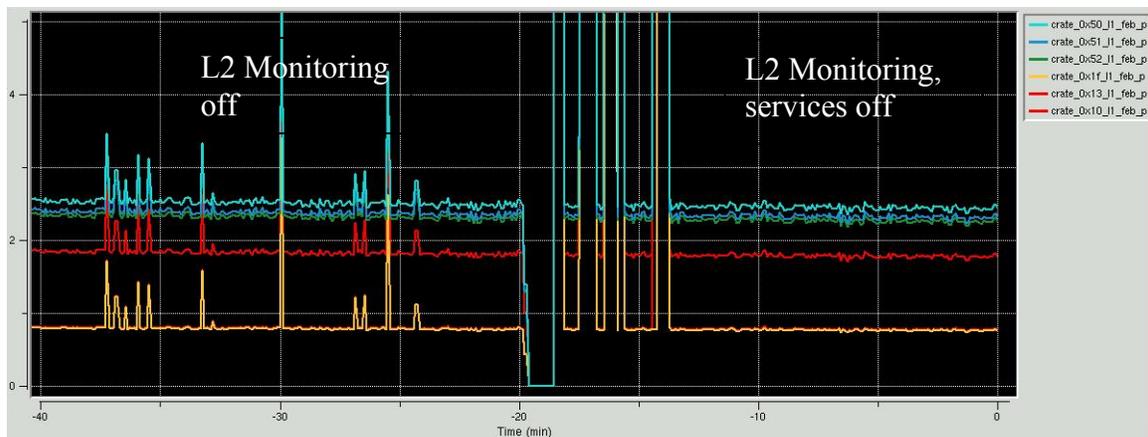


Figure 9. Stripmon display of FEB with L2 monitoring off and L2 monitoring and services off.

running at monitoring level 0, we now run at monitoring level 1 (20% faster than level 0) which provides information on occupancies and sync errors necessary to debug missing

inputs. With a day of work a new (0.5) monitoring level can be implemented which should reduce the monitoring dead time by a further 50%.

Level 2 monitoring currently runs asynchronously in each crate. So the deadtime due to monitoring adds in a semi-random fashion. Level 2 has the ability to use the “collect status” modifier bit to order all of the crates to collect monitoring information simultaneously. The simultaneous monitoring of all crates using the collect status bit has the capability of proving the deadtime due to L2 monitoring by ~30%. This is currently being implemented as a special L1 trigger which fires every 5 seconds and is always rejected by level 2. A proper hardware implementation of this will have to wait until the summer shutdown.

Level 2 services also cause an observable amount of FEB. These effects have also been seen in timing studies of the level 2 beta computers. It was decided to keep services at the current level for the alphas and work on improvements for the betas, which will be installed in the next two months.

5 Summary

We now believe that we understand the sources of deadtime in considerable detail. There are ongoing studies to understand how to minimize both the fixed fraction from Level 2, the event overhead in the SMT and CFT, and the data load in the SMT. Incremental gains can be achieved by improving SMT monitoring and thresholds, a better understanding of the systematics of tick dependent pedestal shifts and coherent noise, and trimming of SVX sequencer timings.

This note has not addressed several related issues. The data load in the CFT is known to be primarily physics-induced. We have not studied the systematics of the data load as a function of luminosity or the effect of changing CFT thresholds on the data volume and tracking efficiency. There have also been studies of the tick-dependent pedestal shifts in the SMT and CFT which can also effect data load and deadtime. These will be addressed in a separate note.

At the conclusion of these studies we can imagine the following scenario:

- CFT – remove VSVX readout, trim sequencer timings, remove unused channel readout
- SMT – trim sequencer timings, install 300 hit cutoff
- Level 2 – Reduce monitoring detail and rate to lower the associated fixed dead time to less than 0.2%

Given these changes the resulting L1 accept rate limitations for the SMT and CFT (assuming no L2 monitoring) shown in table 3 are equal within errors.

	Overhead (μ s)			<Hits>			L1 Rate		
	Now	2 Mo	Final	Now	2 Mo	Final	Now	2 Mo	Final
CFT	17.1	13	9.3	225	225	225	1500	2000	2662
SMT	14.5	9.3	9.3	421	421	272	1600	1900	2511

Table 3. L1 rates at 5% front end busy now, assuming improvements in SVX sequencing (2 mo), and after all suggested changes are implemented.

The “now” rates are faster than we can currently run because the L2 monitoring deadtime has not been included. It is unlikely that these rates can be improved significantly in Run 2a without major changes in the electronics or data acquisition.