Workshop on
High Performance, Fault Adaptive
Large Scale Real-Time Systems
Vanderbilt University Nov. 14–15, 2002

High Energy Physics Triggers and Data Acquisition Systems

Peter S. Cooper
Fermi National Accelerator Laboratory

Abstract

I will review the nature and challenges of real−time data selection (triggering) and data acquisition in modern High Energy Physics experiments using examples from my recent and future experiments.

Outline

I'll first go thru SELEX, our recent charm experiment at Fermilab, as a case study of Triggering and DAQ in HEP. Then I'll look at future experimental requirements and challenges with examples from CKM (you've heard about BteV yesterday from Joel Butler).

- High Energy Physics
  - Goals
  - SELEX − a recurring example
  - Experimental environment
  - Trigger / DAQ system example
  - Some concepts and definitions (*a la* HEP)
  - Faults and their relative tolerability
- Future challenges
- Conclusions

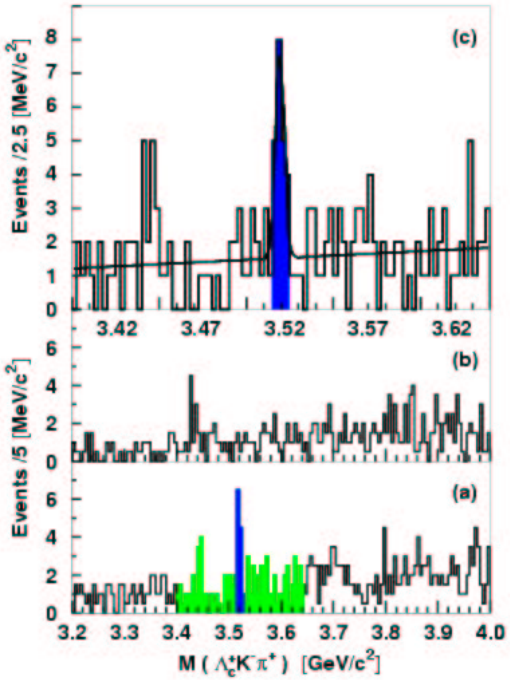
HEP Physics Goals

The study of the states of matter and their interactions at the smallest scales.

A recent new result from SELEX

- A new state in the neutron's family
- As heavy as a Helium atom
- Decays weakly (like the free neutron)
  - Much shorter lived (<30 fsec / 886 s)
  - ~1 mm in the lab (γ~60)
  - Complicated multi−particle decay chain

    $\Xi_{cc}^{+} \rightarrow K^{-}\pi^{+}\Lambda_{c}^{+} \rightarrow pK^{-}\pi^{+}$
- ~16 events in a recorded sample of 1×10^{9} filtered triggers from 15×10^{9} interactions
- All particles are precision tracked, identified, analyzed and reconstructed.
- This event was over in 200 nsec!



HEP Experimental Environment

In HEP we have to first create the states of matter under study. In a time measured typically in nanoseconds the particles created have decayed and / or left our detectors. In an earlier era we recorded events by taking pictures, in a bubble chamber, of the ionization trails left behind. Now we build electronic bubble chambers which produce very large volumes of data.

Our data are intrinsically statistical in nature. Each event is completely independent of all the others. The correlations among events are the physical laws we seek to study. Losing a small fraction of the events is the price of doing business (deadtime). Losing, or garbling the data within one event, if undetected, is potentially deadly.

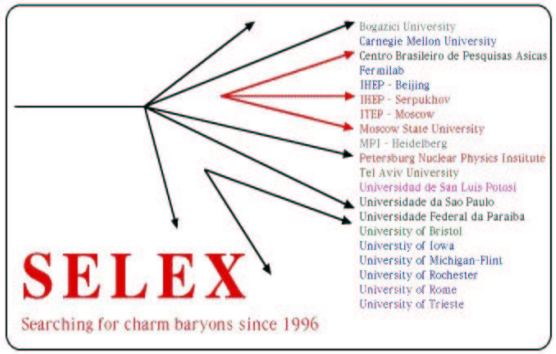
"Smashing atoms" is an expensive sport. Fermilab's budget is ~300M$ /year, much of which is to operate the accelerator complex. Experiments must be up and "alive" when beam is available. When the run is over you're done. There is a huge premium on experiments which work right the first time (or at least sooner rather than latter).

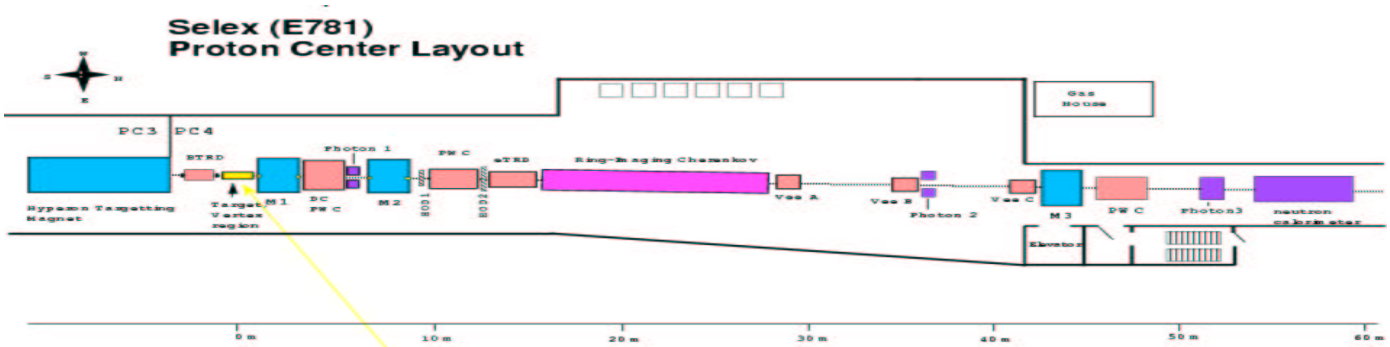
Getting the right data selected and recorded with minimum loss while knowing that the apparatus is functioning properly is the challenge.

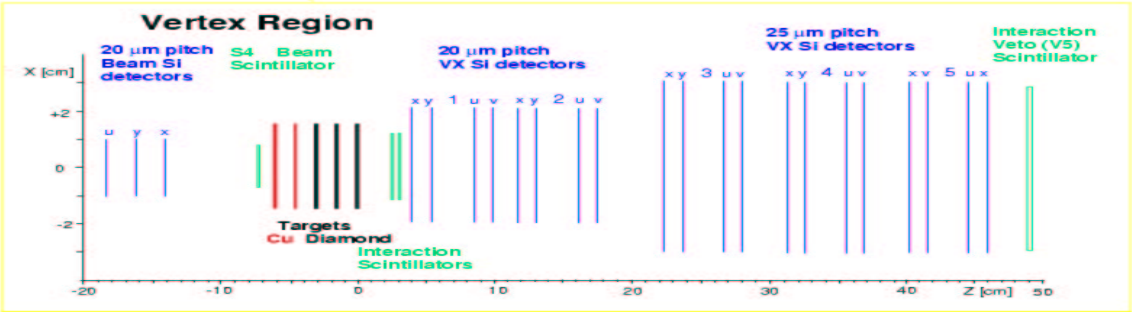
The SELEX Experiment [1996- 7]

Fixed target – charm hadroproduction and decay

- 125 Physicists / 20 Institutuions / 11 countries
- ~100,000 readout channels
- 12−5 separate detector sub−systems
- ~500 kHz beam, ~10 kHz interactions
- ~5 kHz Trigger rate
- ~6.5KB /event sparsified fi 33 Mb/sec
- Everything is in a radiation area 100m from us.







SELEX's DAQ Strategy

- In SELEX we wanted to produce and study particles with charmed quarks. The probability for a charmed particle to be produced, decay in a particular mode and to detect all the particles in the decay mode is:
$\sim 10^{-3} \times 10^{-2} \times 10^{-1} = 10^{-6}$.

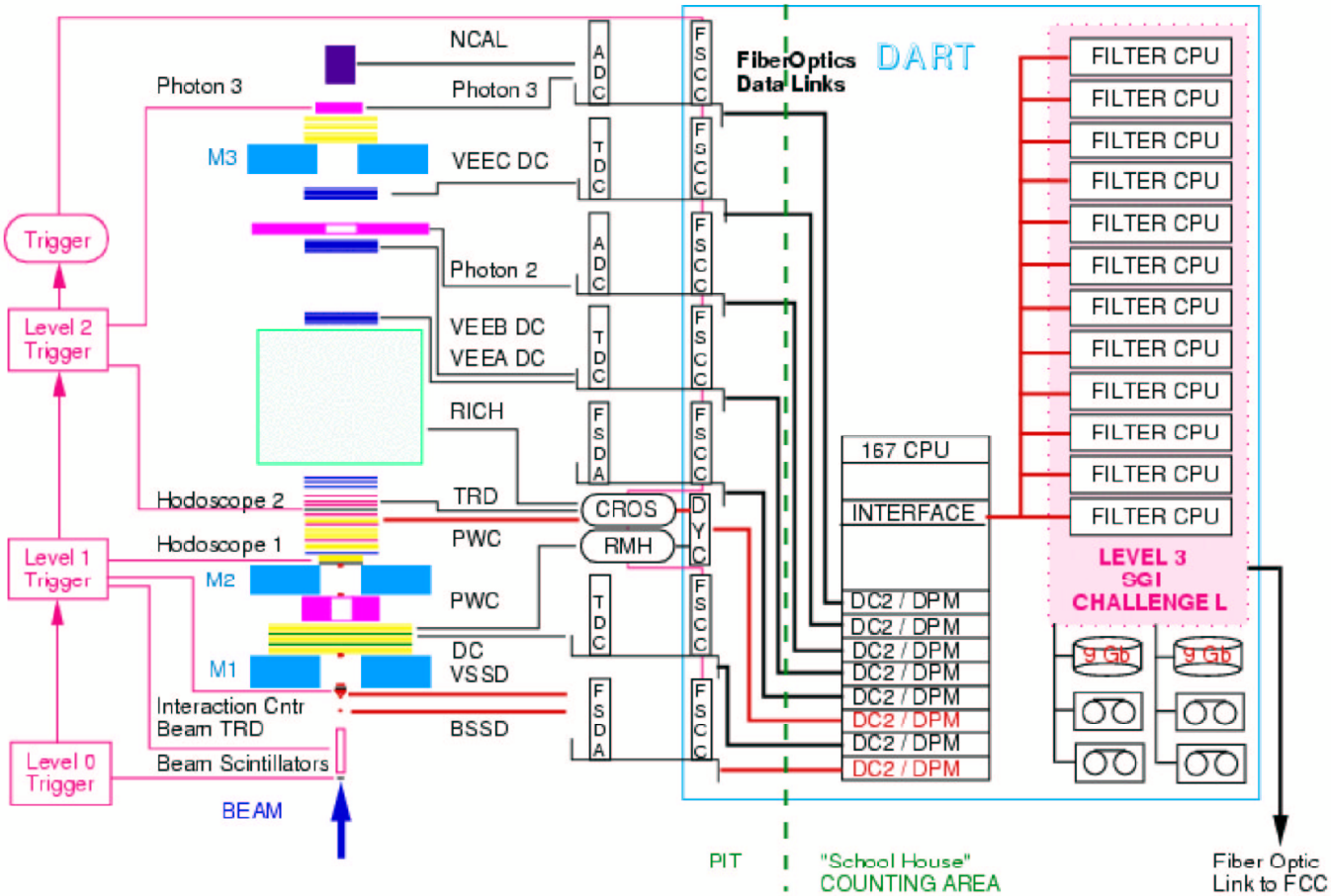
- With electronic signals from the detectors we made a very simple Trigger, very near the apparatus, which required an incoming beam particle and an interaction occurring in our charm targets. Triggered events were digitized, incurring deadtime, and written to shared memories via parallel data streams.

- Each trigger's data were filtered by a farm of 24 unix processors. Since charmed particle fly a small distance (~5mm) before decaying the filter algorithm was to require all reconstructed tracks to be inconsistent with a single vertex in space. Triggers which pass the filter are built as events and written to storage (8mm magtape).

- The filter reduced the data volume by a factor of 8 with an efficiency of 50% enhancing the charm signal by a factor of 4.

psc 11/29/95

E781 TRIGGER / DAQ / ONLINE FILTER SYSTEMS



Some concepts and definitions (*a la* HEP)

- Trigger — A selection done early, usually with hardware and usually on raw, or even analog signals. Triggers have levels, you can pass T1 and fail T2. Triggers are fast and rigid.
- Digitized — The conversion of analog electronic detector signals into digital data. This is usually sparsified data with all non−active channels suppressed.
- Deadtime — The time, or fraction of the time when the DAQ cannot accept new data, usually because it is busy with present data (more below).
- Data streams — To go faster modern systems run parallel DAQ's on sub−systems of the detector. This requires fanning out triggers and building events.
- Filter — A selection done after the trigger, usually with software. These can be quite complicated − up to complete reconstruction of the event.
- Farm — A dedicated set of processors and associated networks or other IO to run a filter, an reconstruction, or a simulation in parallel (event independence!).

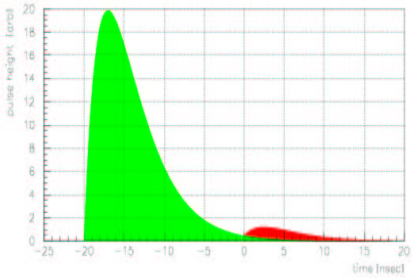
Some concepts and definitions (a la HEP) [continued]

- Reconstruction — The codes to convert a bag of sparsified hardware bytes into physics objects (tracks, vertices,4−vectors,...)
- Event building — The process of catenating event fragments from all data streams into an event.
- Efficiency — The probability to keep the signal e_s or (background e_b)
- Enhancement — $E = e_s / (e_s + e_b)$ The "enrichment" of the sample after a selection

Deadtime − the critical concept

- Deadtime is the loss of acquirable events. It occurs on all timescales:
  - 10^7 s The run is over − you have all the data you will get .
  - 10^5 s An accelerator component fails − lose a day.
  - 10^4 s Access to fix a failed detector component − lose some hours
  - 10^3 s The DAQ crashes and needs a cold−start (reboot−all?)
  - 1 ms A FIFO in the DAQ fills and asserts busy (inhibits triggers)
  - 10 μs An old style (Wilkenson) ADC is dead while digitizing.
  - 1 ns Two interactions happen at "same" time in the apparatus.

- Contributions to the deadtime fraction go as $R_i t_i$ where R is the rate of a problem which incurs a deadtime t .



•Some deadtime contributions are out of the experimenter's control. These set the scale for all others.

•A reasonable goal is to keep controllable contributions to <1%.

Data Quality − Deadtime in disguise

- There are many cases where data acquired while the experiment is apparently alive must be discarded due to detector, front−end electronics or DAQ problems. Identifying and correcting these problem ASAP is critical.

- Monitoring the data and detecting and reporting potential problems in heterogeneous distributed systems like these is difficult.

- These problem are often repaired by reloading / restarting a low level subsystem. The trick is to identify which sub−system and to have engineered the whole DAQ so that partial restarts are possible. I've never seen this successfully done in a systematic way. The time to restart a DAQ is an important parameter which is always out of control.

- Much of the physicist time on "shift" taking data is spent in this area.

- We ought to be able to do better with better planning, engineering and more modern technologies.

Faults and their relative tolerability – some examples

- Data corruption (e.g. building events with fragments from different triggers) is unacceptable. Systems should be designed to forbid this and should slam on the brakes if it is ever detected.

- Data transmission error detection and corrections a delicate design balance. Retransmissions can cause more, or less, deadtime than just flagging 0.1% of the event as bad and flushing them.

- Data sources to the DAQ must be alive and known to be alive. You can't reboot the data source of a critical detector during a data taking! Data sinks (e.g. farm nodes) can and will go to lunch from time to time (crashes, loops too much data, a really long event,...). Modern systems must cope with this. Recovering 1 node in 100 during a run is probably a bridge too far.

- Analyzing and simulating these complicated heterogeneous system to make intelligent design choices is a non−trivial undertaking

Future Challenges

- Yesterday, Joel Butler discussed BTeV and it's not inconsiderable goals.

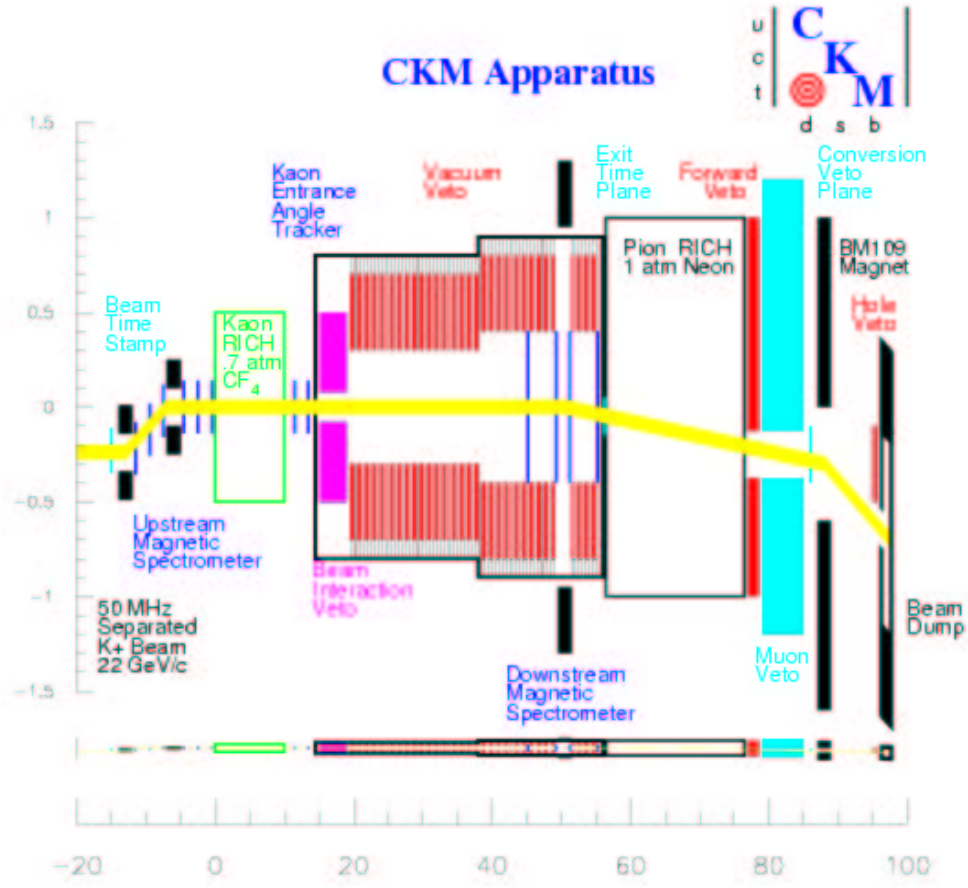
- Our new experiment is CKM, the measurement of the rate of the decay mode K^+ fi p^+n' n. The occurs in 10^{-10} of decays! There is a background process (K^+ fi p^+p^0fi gg) which happen 1 billion times more frequently.

- This is a high rate experiment (50MHz beam, 10MHz decay rate) with exceptionally low inefficiency requirements (10^{-5}/m, $3x10^{-5}$/g, 10^{-7}/p^0). These inefficiencies must include all undetectable losses.

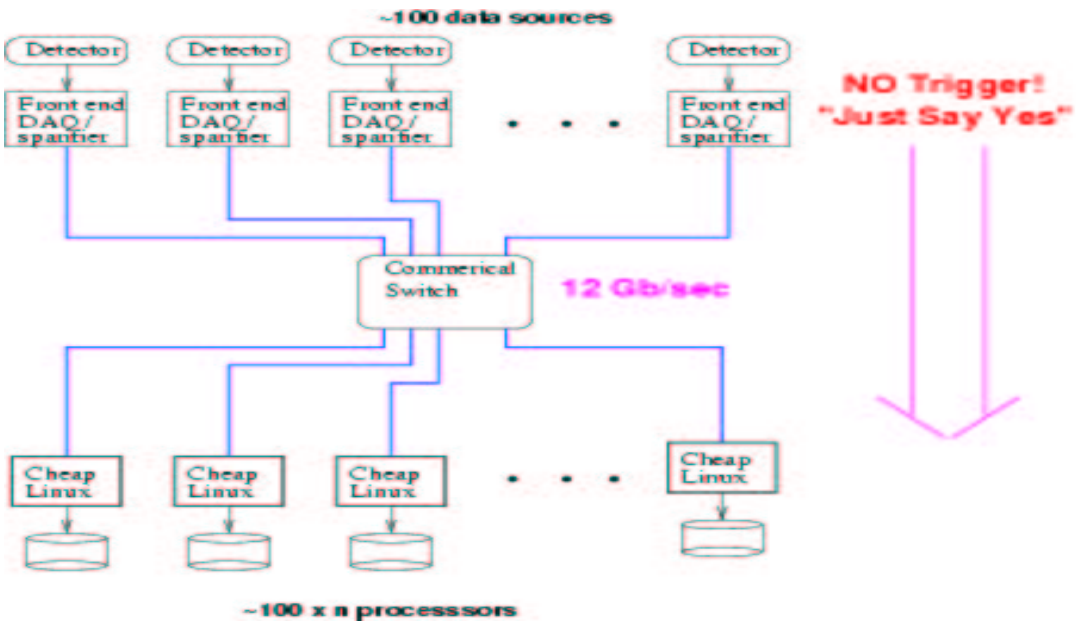
- The levels of performance, and our understanding thereof, required of all aspects of this detector, including the electronics and DAQ, far exceeds the usual HEP industry standards.

CKM Experiment

- A decay in flight apparatus
- One and only one charged particle at each z
- Veto everything else
- Redundant independent detectors for each measurement and veto − inefficiencies must be measured.
- Long (110m), appartus makes distribution of timing, data and triggers hard.
- In order to understand deadtime in the veto systems each events is a time interval of 100−200 nsec. Channels can he hit multiple times.



CKM DAQ Strategy



- NO trigger! "Just say Yes".
- Send all sparsified data to a farm via a commercial switch
- Emulate trigger is software in the farm. Maximum flexability.
- All data are times or keyed by time. "Event" is a time interval.
- Concept now in simulation.
    - Can we make the trigger with ~100 processor and not ~1000?
    - What are the real network requirements?
        - error rates / recovery
        - achievable bandwidth
- Plan B − some, not all of the data is trigger farmed.

Bowden's Theorem −If you can't buy it at Circuit City its not commercial!

Conclusions

- HEP Trigger and DAQ systems have been at the (b)leading edge of network and computing technology for many years. (We used to design and build our own microprocessors and networks.) Things have only gotten easier with the advent of commercial components and serious standards (ethernet, TCP/IP...)

- Computing is still enabling technology for HEP. We will push it as hard as we can, and probably harder than we should, because it allows use to do experiments which would not otherwise be possible.

- Experiments with 1000's of realtime networked processor are on the books.

- We should satisfy my 1st theorem for experiments "Make no old mistakes." We will, however, find new mistake to make. Just imaging 1000's of nodes all crashing at once; consider the reboot over the network!

- Serious expertise in building systems like these which really work reliably is of great value to these projects.