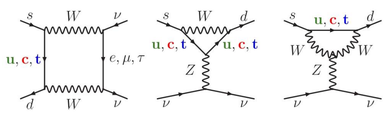
GPUs for fast triggering and pattern matching at the CERN experiment NA62

G.Collazuol [1], G.Lamanna[1,2], M.Sozzi[1,3]

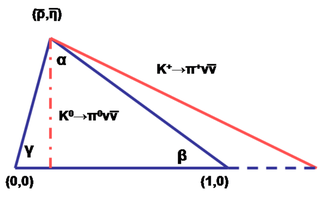
([1] INFN sez.Pisa, [2] Scuola Normale Superiore of Pisa, [3] University of Pisa)

From theory to experiment

- $K^+ \to \pi^+ \nu\bar{\nu}$ process is quite unique as extremely accurate and clean probe for the non-trivial flavour structure of physics beyond the Standard Model



- $K^+ \to \pi^+ \nu\bar{\nu}$ BR (Branching Ratio) is computed in a clean theoretical environment due to the small contribution by hadronic matrix elements and long distance terms.

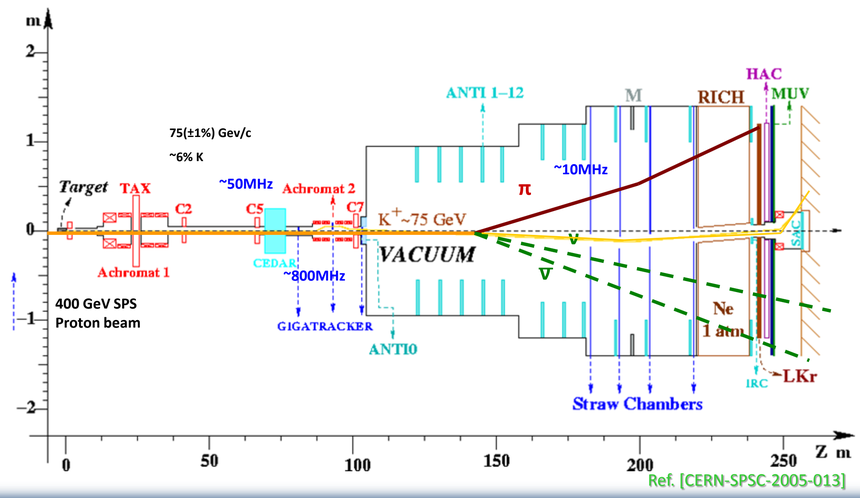


- FCNC process forbidden at tree level (GIM meccanism)
- Top contribution is dominant in loops: cleanest way to extract V_{td} and to give independent determination of the unitarity triangle

- **NA62** aims at measuring $K^+ \to \pi^+ \nu\bar{\nu}$ with O(100) SM events, in two years of data taking using novel decay-in-flight approach
- Very challenging experiment: veto capability on not-interesting decays and high beam intensity
- tracking of single beam and decay particles, redundant time, momentum, angle measurements, high vetoing power for charged and neutral particles, highly effective particle identification, ultra high rate capability…

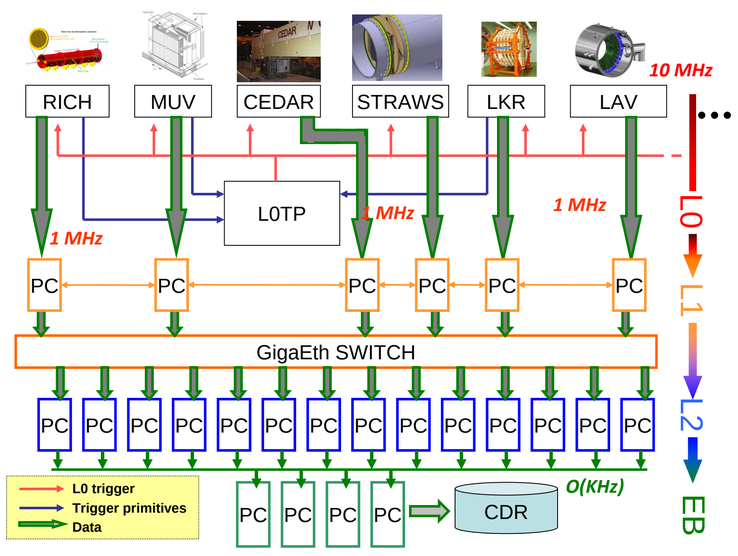
$BR(K^+ \to \pi^+ \pi^0)$	$20.7 \cdot 10^{-2}$
$BR(K^+ \to \mu^+ \nu)$	$63.5 \cdot 10^{-2}$
$BR(K^+ \to e^+ \pi^0 \nu)$	$5.1 \cdot 10^{-2}$
$BR(K^+ \to \mu^+ \pi^0 \nu)$	$3.4 \cdot 10^{-2}$

$$BR(K^+ \to \pi^+ \nu\bar{\nu}) = (8.5 \pm 0.7) \cdot 10^{-11}$$



Ref. [CERN-SPSC-2005-013]

The NA62 Trigger

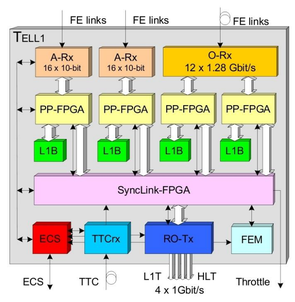


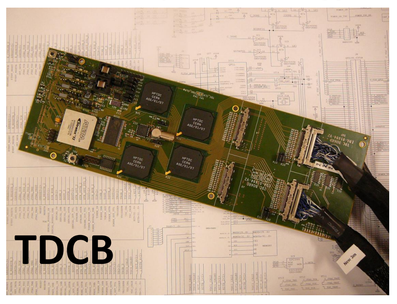
- Multilevel trigger: high flexibility
- L0 hardware, L1 and L2 software
- The rate will be reduced from 10 MHz to few tens KHz
- Integrated trigger and DAQ in a fully digital system
- Fully monitored system: undetected acquisition losses $<10^{-8}$
- L0 central processor (L0TP) L0 latency <1 ms (large buffers)

detector	Rate (MHz)
CEDAR	50
GTK	800
LAV (total)	9.5
STRAW (each)	8
RICH	8.6
LKR	10.5
MUV	9.2
SAC	1.5

The L0 bulding blocks

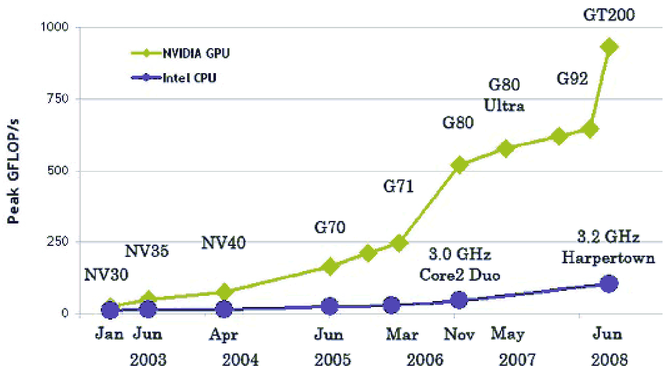
- TELL1:generic DAQ board (EPFL Lausanne for LHCb exp.) used both for acquisition and generation of trigger primitives
- New TELL1, with more powerfull FPGAs and larger buffers, has been by the NA62 Pisa Group
- Logic, primitives definitions, monitoring and buffering performed and controlled by 9 FPGAs for each board (5 on TELL1 and 4 onTDCBs) + 350 MB DDR RAM
- Up to 4 custom daughter boards
- TELL1 provides asynchronous GbE data transmission (4Gb/s) and receives synchronous trigger requests via TTC standard





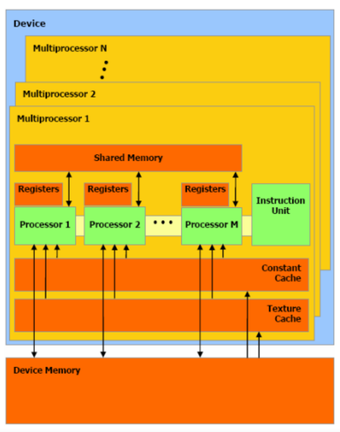
- Analog signals digitized by four 128-channels HPTDC boards (TDCB): 512 channels on a single TELL1 board
- 128 TDC channels per board with 100 ps time resolution
- QPLL on board to decrease the clock time jitter < 15ps
- LVDS inputs
- I2C and JTAG used for configuration
- Altera Stratix II FPGA for data monitoring and preprocessing
- 1MB SRAM
- Low noise DC-DC converters and filters
- Miniaturized connectors

GPU (Graphics Processor Unit)



- Fast expanding field of scientific computation on GPUs (GPGPU)
- This project: attempt to exploit GPUs in moderately large hard REAL-TIME environment
- So far large latencies restricted this to high-performance computing; now O(100us) latencies can be obtained

- Easy and cheap on a workstation
  - a few TFlops → fast answers
  - hundreds GB/s → high data rates
- The GPU sector is one of the most supported by IT industry (Digital imaging, video processing, Video Games,…): *can profit for free!*



- GPU = graphics processor in commercial Graphics Cards
- GPUs power increasing faster than the CPUs due to:

(1) Different architecture: more transistors for processing rather than caching, resource usage prediction and flow control

(2) Different scalability: the GPU power is related to the size and number of chips (more and more cores) and not to the clock speed (as in the CPU)

- GPU operate simultaneous processing on thousands points with a multilevel SIMD (Single Instruction Multiple Data) architecture: large parallel computing power, large memory, huge bandwidths
- Examples:

ATI-RADEON 4870:
10cores x16processors x 5sub-units
~1TFlops, ~100 GB/s bandwidth

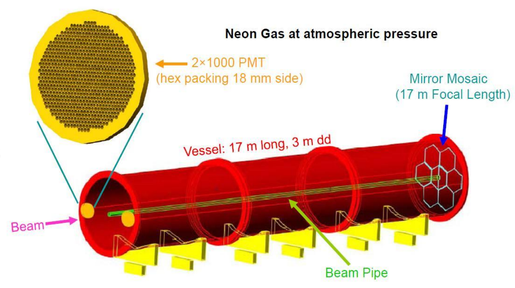
NVIDIA-TESLA C1060:
30 cores x 8 processors + 4GB RAM
~1TFlops, ~100 GB/s bandwidth

GPUs	1 Tesla GPU
Single Precision Performance	933 Gigaflops
Double Precision Performance	78 Gigaflops
Memory	4 GB DDR3
Memory speed	800 MHz
Bandwidth	102 GB/s

RICH & Rings

- RICH counter (Ne, 1atm) for particle identification downstream the decay region.
- separation $\pi-\mu$ at $5x10^{-3}$ level (range 15-35 GeV/c)
- Good time resolution: below 100 ps
- ~2000 photomultipliers to identify Cherenkov rings
- Partecipate to L0 trigger definition with multiplicity (in the default implementation)

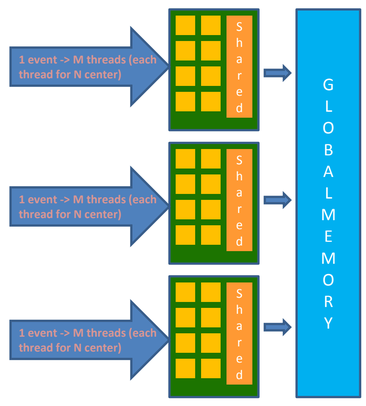
- Rings at trigger level → selective conditions
- Ring's radius and center → velocity and angle → coarse on-line measurement of particle momentum (with assumption on particle type)

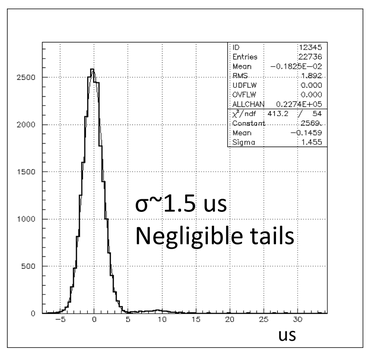


- Ring finding at 10 MHz is a challenging task
- On average one pion results in 20 hits → 1.6 GB/s (160 B/event)

Fast Ring finder

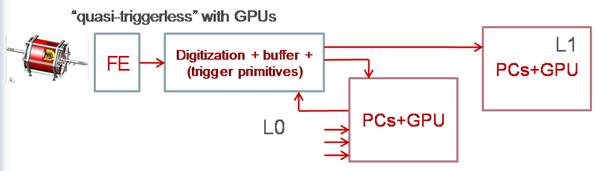
- Several ring finder algorithms studied. More suitable for the parallelization: Generalized Hough Transform and Multi-Histograms of distances
- In order to obtain the best performances it's important to correctly exploit the different types of memory available and the architecture of the device
- The parallelization is employed at two levels: (1) the algorithm is parallelized to speed up the execution on the single event and (2) many events are processed at the same time





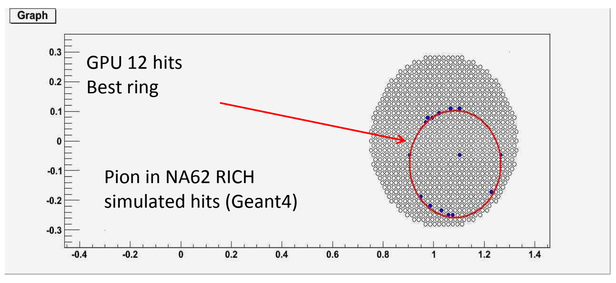
- The data transfer time between PC and Video Card, determined by the PCI Express gen.2 bandwidth, is optimized by sending large packets of events concurrently with the execution
- The effective time to find a ring, in 1000 events per packet, is: 3.1 μs/event in the best obtimized approach
- Stable timing results with negligible tails
- "On track" for use in <1ms latency on-line system
- Further improvement is expected from additional optimization, new algorithms under study and more powerfull devices already available on the market (for instance NVIDIA FERMI Video Card will provide 2 Tflops with double bandwidth)

GPU integration in NA62 trigger



- Two possible applications: in the software levels (L1 and L2) and/or in the hardware level (L0)
  - → L1 (L2) : no latency requirements, useful to decrease the dimension of the L1 (L2) farm
  - → L0: latency < 1ms, effective and selective trigger decisions
- Simpler scheme for GPU cooperation with L0TP: Gigabit connection with R/O cards and links through PCI-e lanes
- Faster approach: connection with special Gigabit card for direct DMA with the Video Card memory and real-time operating system on PC (under study)
- Many of the detectors can participate to lowest trigger level thanks to the GPU approach (spectrometer, calorimeter, etc...)

- Measurements with high resolution at L0 = great opportunity for efficient trigger in ultra high rate environment
- Total trigger bandwidth for main physics channels (no zero-suppression) can be reduced by the use of selective conditions early in the trigger chain
- TDAQ will not be a bottleneck in case a higher beam intensity is used
- The scheme will give the unique possibility to collect additional interesting decay channels, otherwise suppresed by the main trigger
- Computing power pressure on the higher software trigger levels highly reduced



System used for tests: PC with
- CPU Intel i7-950 +12GB RAM,
- GPU: 2 x NVIDIA TESLA C1060
i.e. 2 x {240 cores + 4 GB VRAM}

gianmaria.collazuol@cern.ch, gianluca.lamanna@cern.ch, marco.sozzi@cern.ch