A novel continuous readout for the NA62 data acquisition system

Marco Boretto, Enrico Gamberini CERN

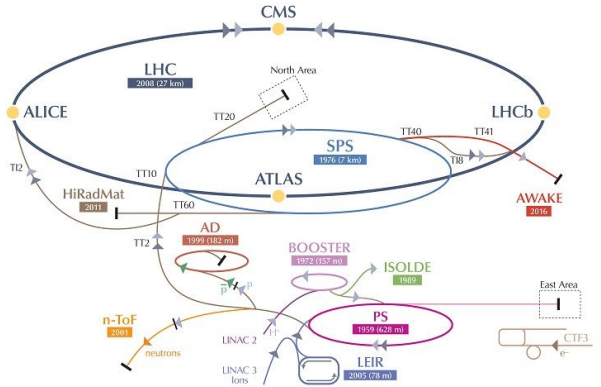
Workshop IX on Streaming Readout - 8 December 2021

The NA62 experiment at CERN

NA62 is located at the CERN Super Proton Synchrotron

Aims at measuring the extremely rare kaon decay:
$K^+ \to \pi^+ \nu\nu$ BR($K^+ \to \pi^+ \nu\nu$) = (8.4±1.0) x 10^{-11}

10^{13} kaon decays need to be collected to measure it with a 10% precision



Data-taking	% of nominal intensity(*)	Decays recorded
2016	40%	5 x 10^{11}
2017-2018	60%	8 x 10^{12}
2021	100%	

100% of nominal intensity: kaon rate = 45 MHz

The NA62 trigger system

To reduce the data rate at a manageable level, NA62 uses two levels of triggers:

The **Level-0 (L0):** a hardware trigger implemented in an FPGA board called **L0TP** (Level-0 Trigger Processor).

Simple detectors associations.

Maximum trigger rate output: 1 MHz

The **Level-1 (L1):** a software trigger running on the DAQ-farm.

Complex algorithms.

Events rate reduction: 1 MHz → 100 kHz

Current NA62 data acquisition system

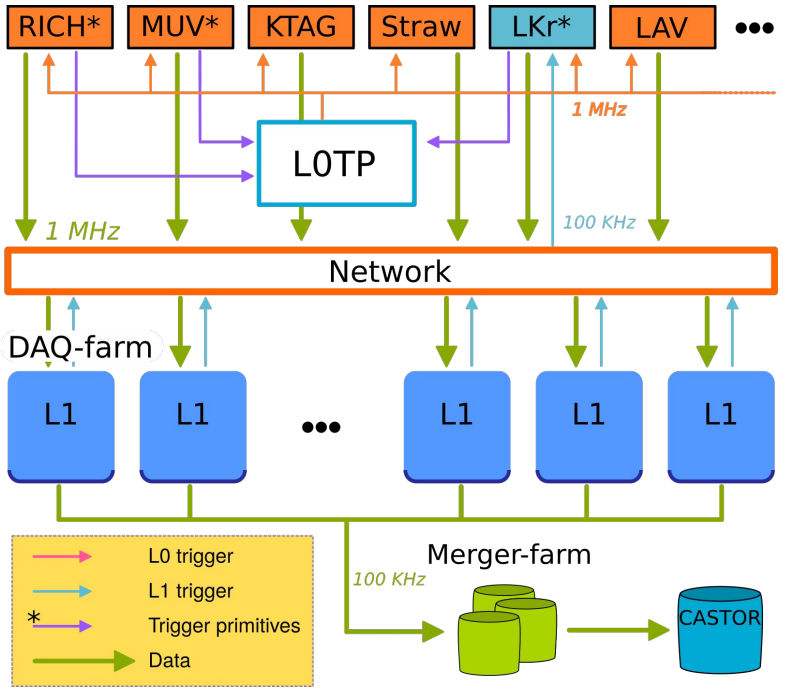
A subset of detectors produce the **trigger primitives**, the input of **L0TP**.

The readout generates the trigger primitives. The computation algorithm must not exceed 100 μs.

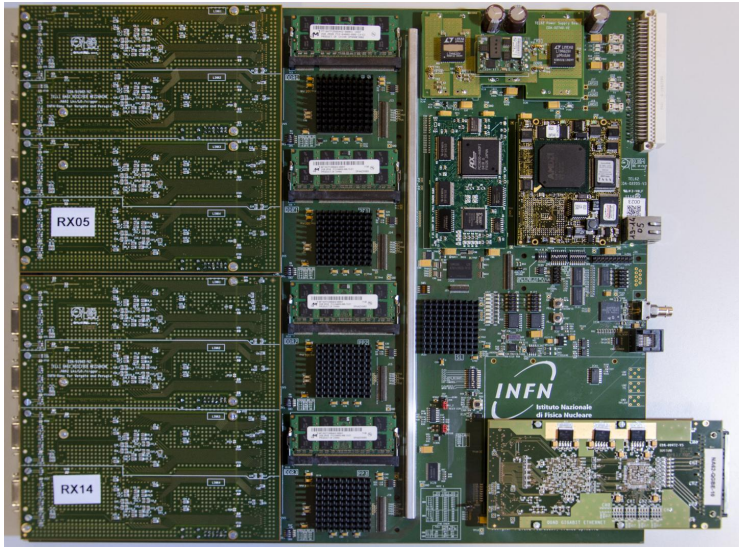
L0TP distributes the L0 triggers via TTC (Timing Trigger and Control system)

A partial event is built in the DAQ-farm.

The detectors that contribute the most in the event size are read out only after a positive L1 decision.



The TEL62 readout board

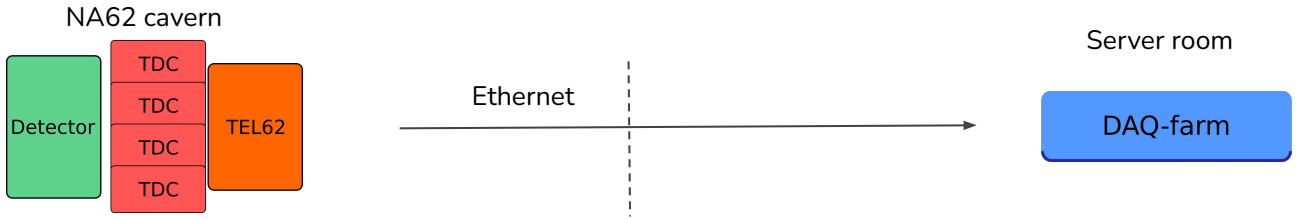


The TEL62 is an **integrated trigger and data acquisition board** used by the most of the detectors.

It can host up to four high-performance TDC boards

- Developed from the TELL1 board for the LHCb experiment (2006)
- Not designed to be radiation tolerant
- Has limitations at nominal intensity especially for high rate detectors
- Limited number of spares available

The TEL62 readout concept



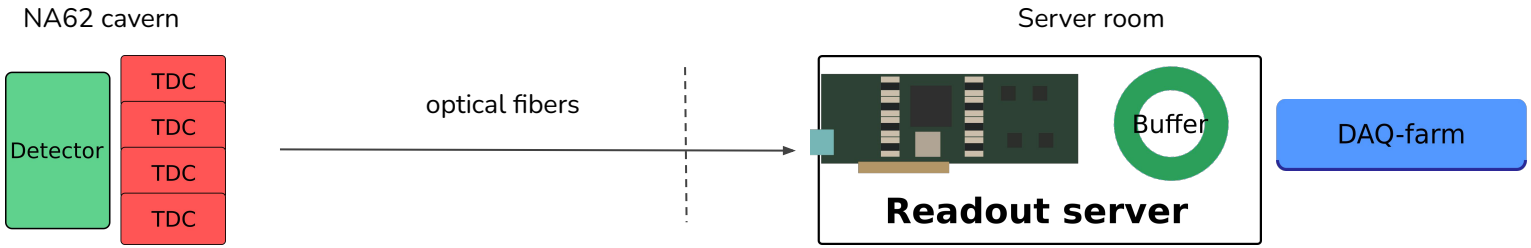
The TEL62 needs to perform lots of operations:

- Time To Digital conversion
- buffer the data (**limited**)
- produce the trigger primitives
- perform trigger matching
- pack the events in UDP data frames and send to the DAQ-farm

Internal buffer limited to the on-board memory.

For that reason the L0 trigger must arrive within **1 ms**.

The new readout concept



On-detector:

- Time To Digital conversion
- send out the data via optical fibers

Off-detector:

- receive the data via optical fibers
- buffer the data (**all the data can be cached**)
- pack the events in UDP data frames and send them to the DAQ-farm

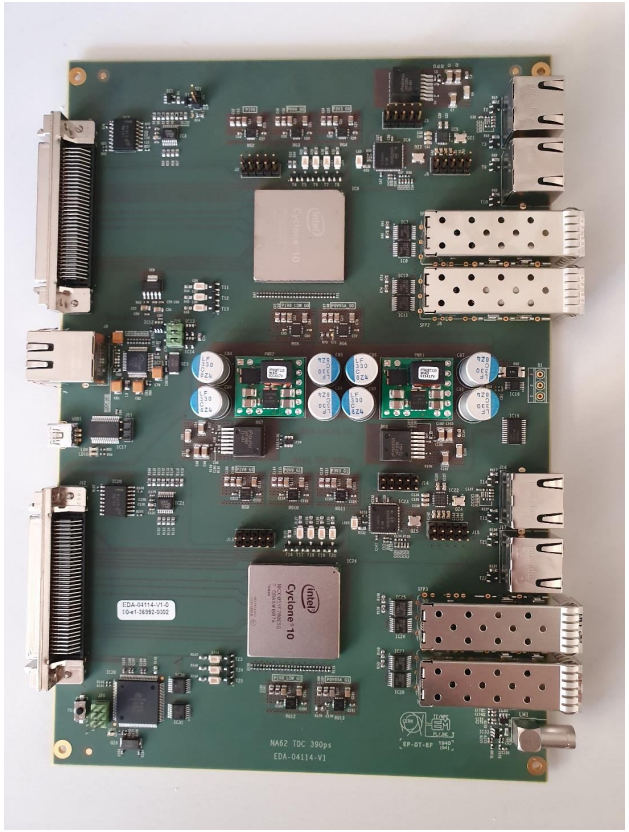
minimise the electronic exposed to radiation

Sitting on a readout server located in a protected environment

On-detector: Custom TDC board

The design includes:

- 2x CYCLONE 10 FPGA (less susceptible to radiation)
- 32 channels with 390 ps bin resolution per FPGA
- 2x SFP+ transceivers per FPGA:
  - 2x 8b/10b encoded TX (FULL mode): data to up-detector and slow control responses
  - 1x GBT RX for TTC and slow control requests



The Front-End Link eXchange board FELIX

The FELIX hardware platform has been developed for the final implementation in the ATLAS Run 3 upgrade.

FELIX is designed to:

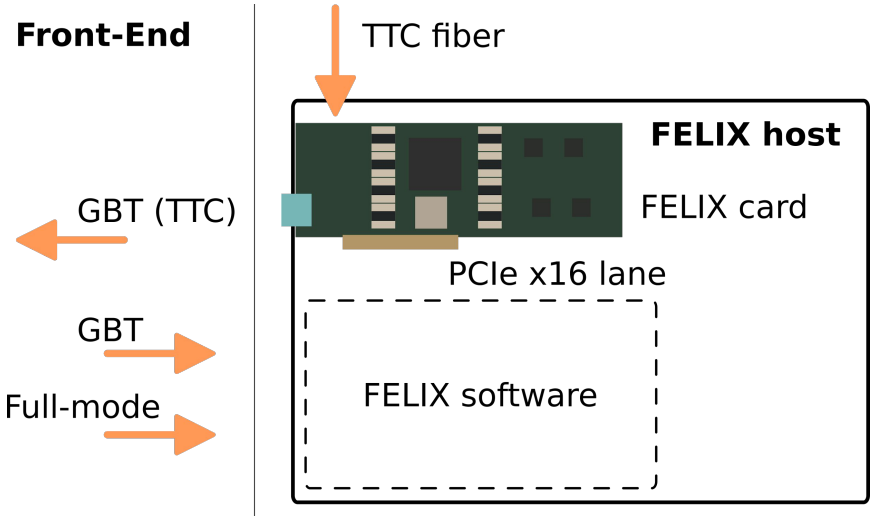
- act as a data router, receiving data from detector front-end electronics and sending on a commodity network through the readout server.
- be detector agnostic.

The FELIX system

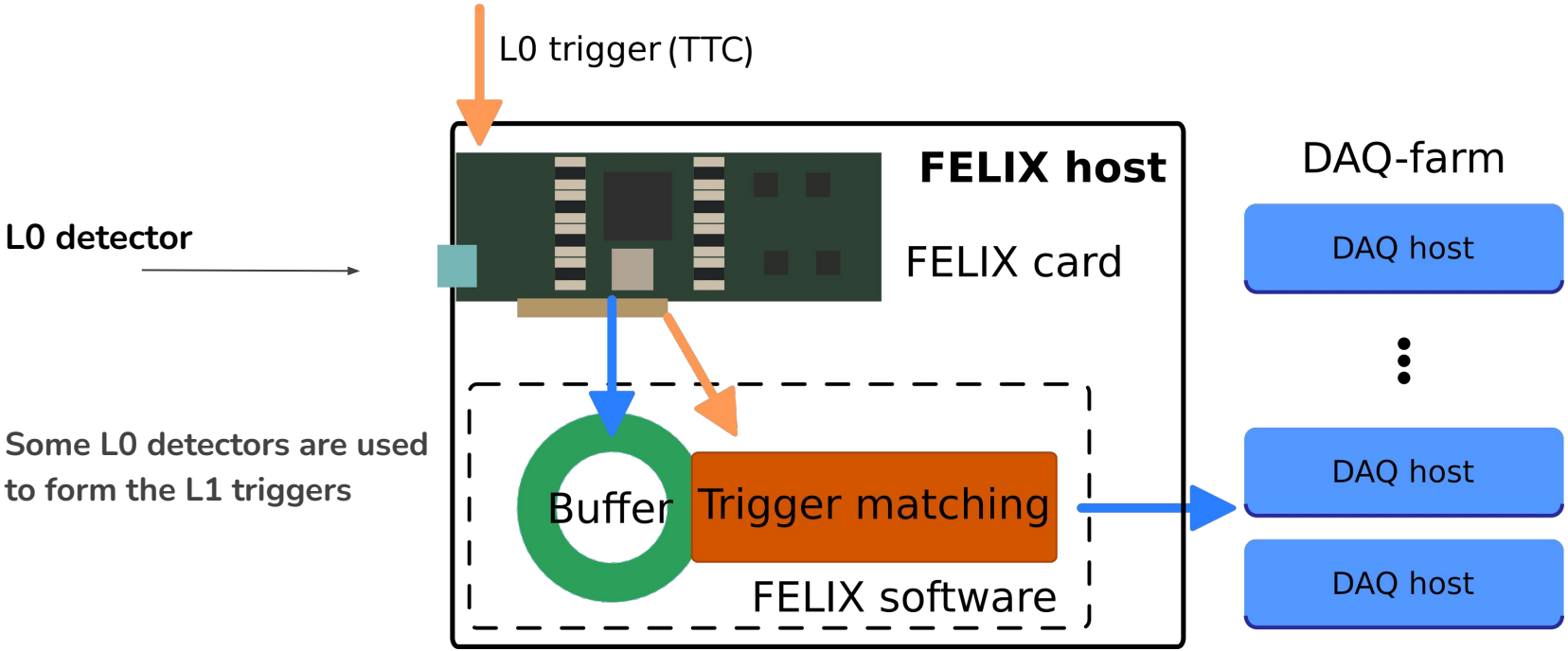
The FELIX system consists of: the PCIe express board, the firmware and the software that run on a Linux server.

Connectivity:
- 2x MTP24 optical link connectors.
- (Optional) TTC mezzanine board.

- FELIX -> Front-End: GBT (GigaBit Transceiver)
  - TTC clock distribution
  - Synchronous command distribution
  - runtime parameter loading
- Front-End -> FELIX
  - Physics data transmission (lightweight FULL mode designed for maximum throughput)
  - monitoring and status data transmission



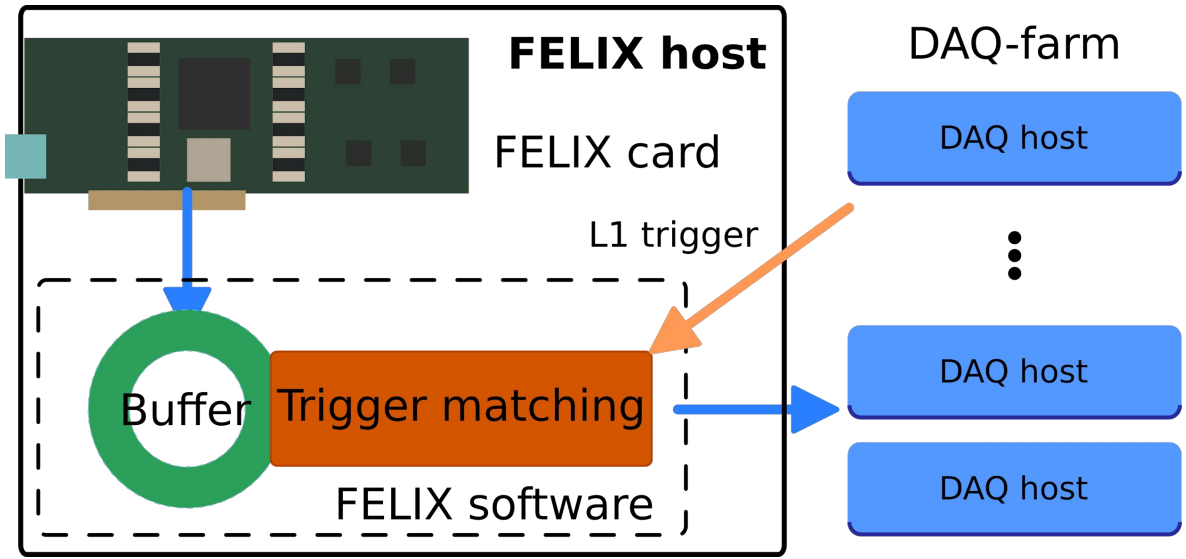
NA62 integration at Level-0



NA62 integration at Level-1

L1 detector

Any detector not used to form L1 trigger verdict can be read out at L1



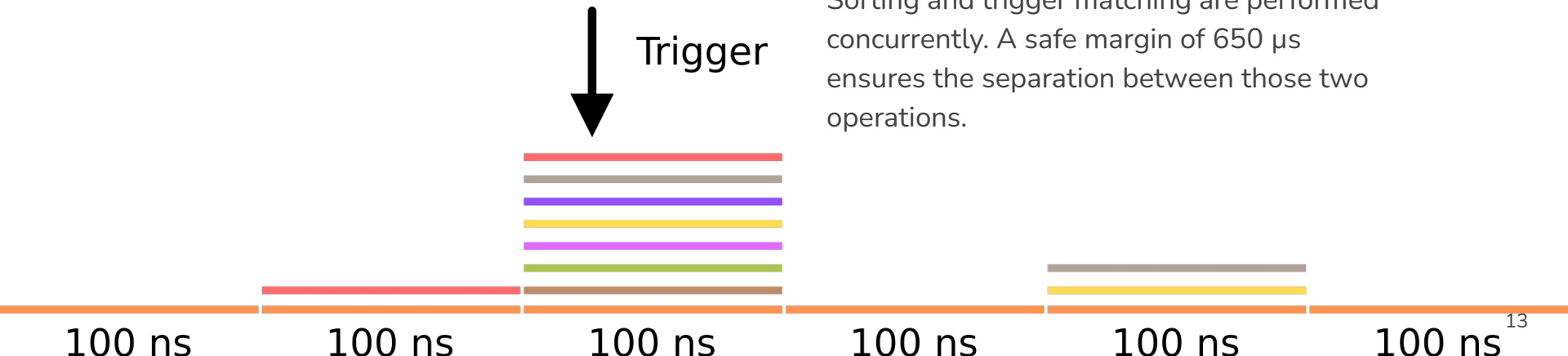
The NA62-FELIX readout software

‘The readout software can be seen as an “in-memory time-based database” were triggers requests are DB queries asking for hits in a time window’

Written in C++, based on the FELIX libraries

- **Sorting:** reads the hits received in the FELIX DMA and sorts them in intervals (e.g. 100 ns)
- **Trigger matching** extracts relevant hits after the reception of the trigger (configurable 25 ns-unit window)

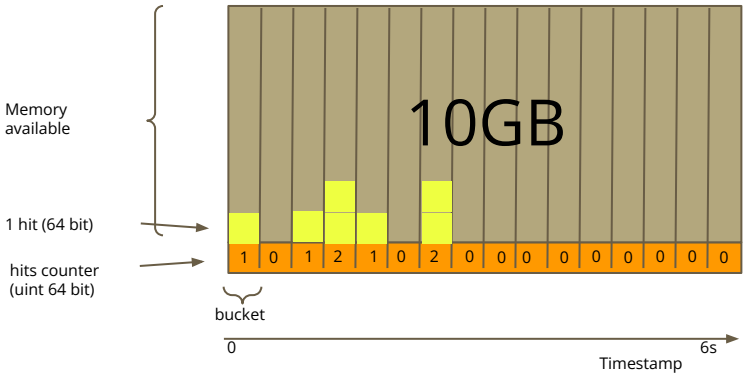
Sorting and trigger matching are performed concurrently. A safe margin of 650 μs ensures the separation between those two operations.



The NA62-FELIX readout software - Continuous Memory

The hits are stored in a continuous chunk of memory, which is divided in two regions:

- **Counters region**: keeps track of the number of hits in the time interval (bucket)
- **Hits region**: host a finite number of hits



- **Continuous chunk allocated as shared memory object**: stays allocated over the lifetime of the DAQ process (fast readout software initialization)
- Size of hits and counter is the same: **memory is perfectly aligned**
- Memory reset: set to 0 the counters (overwrite)

Limitation: can handle a finite number of hits in a bucket

The NA62-FELIX readout software - Fragment Cache

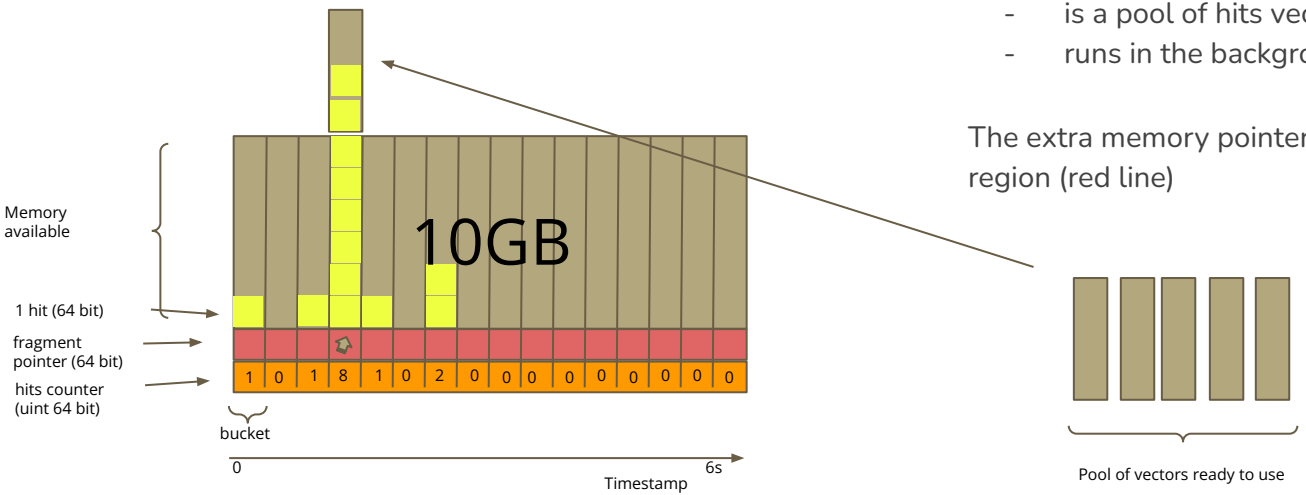
The hits in excess have to be handled by the **Fragment cache** that gives the possibility to dynamically change the size.

The fragment cache works with the continuous memory and provides extra memory whenever a bucket is full.

The fragment cache:

- is initialized at the process startup
- is a pool of hits vector
- runs in the background

The extra memory pointer is stored in the pointer region (red line)



The NA62-FELIX readout software

Using the Continuous memory in combination with the Fragment cache allows to:

- minimise the memory allocations/de-allocations
- handle spikes of hits
- quickly reset the full memory structure

The readout software can

- handle transparently L0 or L1 triggers
- manage up to 12 optical bidirectional links

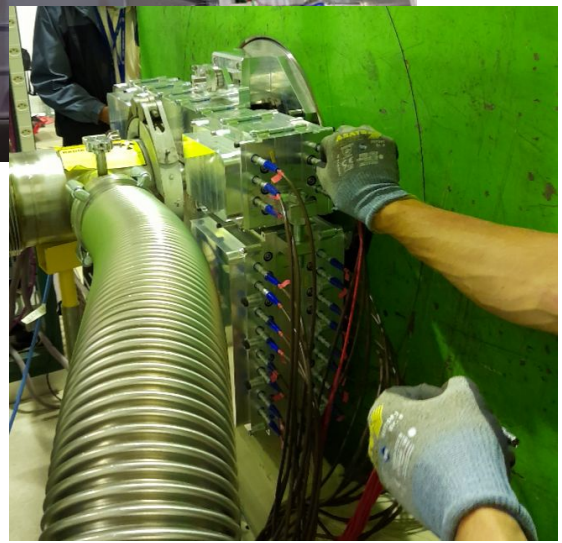
Every optical link is handled in parallel and the extracted triggers are merged and sent out through the NIC.

“Veto Counter” L1 integration

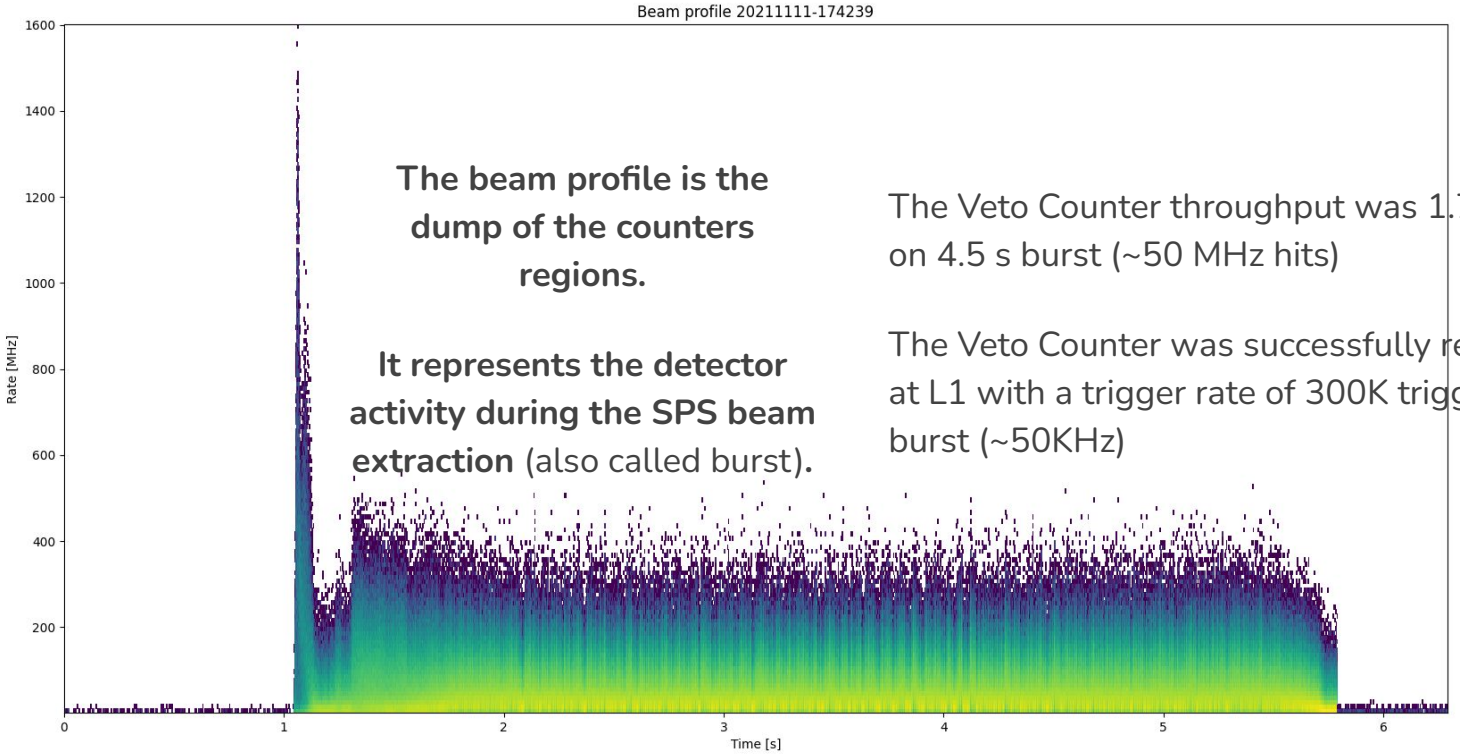
During the 2021 data-taking a new detector, the **Veto Counter**, was developed and commissioned in the NA62 experiment.

The Veto Counter was instrumented with the new readout at L1 as it is not involved in the trigger decision.

64 channels, 2 TDC boards, 2 optical links (only a FPGA per board was used), 1 FELIX server



Veto Counter beam profile



The beam profile is the dump of the counters regions.

It represents the detector activity during the SPS beam extraction (also called burst).

The Veto Counter throughput was 1.7 GB on 4.5 s burst (~50 MHz hits)

The Veto Counter was successfully read out at L1 with a trigger rate of 300K trigger per burst (~50KHz)

“Veto Counter” L1 integration

Veto counter read out with the data acquisition upgrade is an important milestone for NA62:

All of the data produced are cached in memory allowing analysis “on the fly” like the beam profile monitor

The cached data can be also used to generate the L0 primitives relaxing the time constraint.

Veto Counter read out demonstrated the feasibility of the new readout approach and will be integrated on more challenging detectors in terms of hits rate and number of channels.

Conclusions

- The design of the NA62 data acquisition upgrade was presented
- The system has successfully read out Veto Counter at L1
- This is the first step towards a completely asynchronous readout system

What's next?

Read out other two NA62 detectors:

- **CHANTI** at L1 16x TDC boards 32 optical channels
- **KTAG** at L0 9x TDC boards 64 optical channels