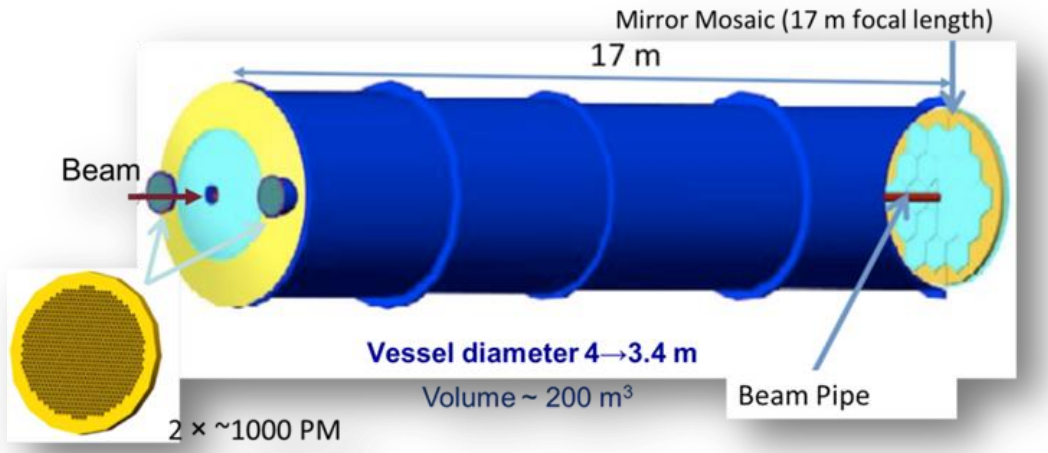
Status report su
GPU-L0TP per il RICH di NA62

Luca Pontisso (INFN – sez. Pisa)
APE Lab (INFN – sez. Roma) for NaNet

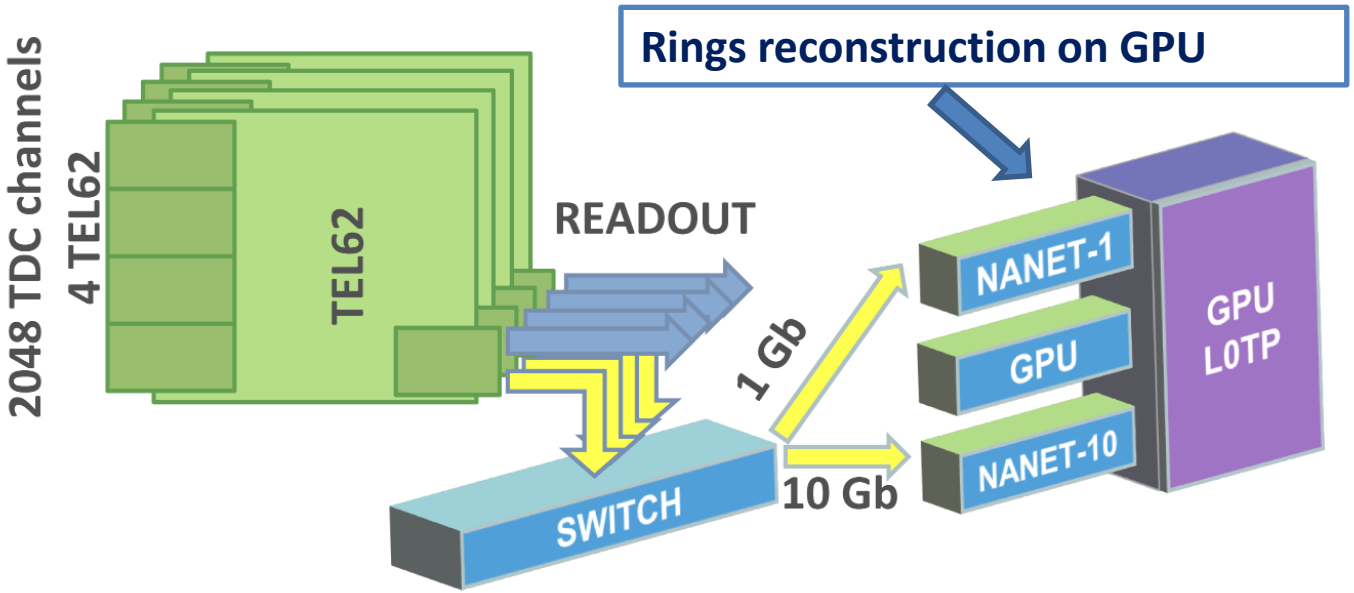
GAP MEETING – 23/03/2016 - Ferrara

GPU-based L0 trigger general scheme



NA62 Rich detector

- Distinguish between pions and muons from 15 to 35 GeV (inefficiency < 1%)
- 2 spots of 1000 PMs each
- 2 read-out boards for each

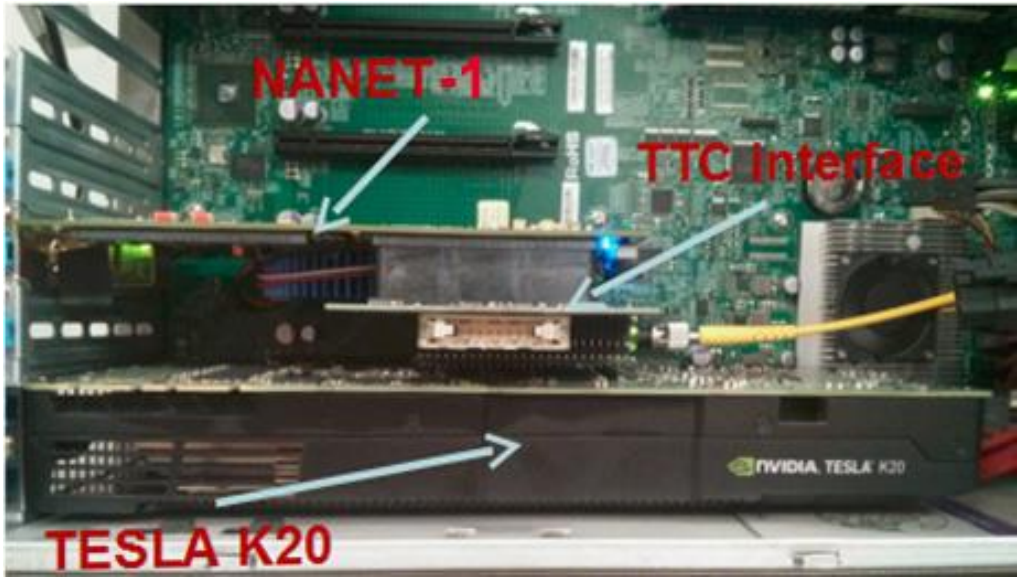


NaNet

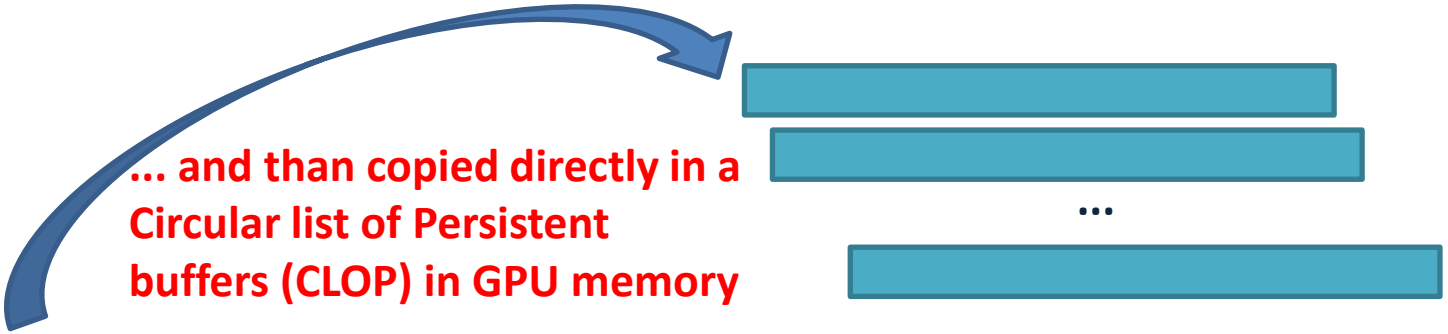
GPU-based L0 trigger setup for NA62 2015 Run

Installed at CERN ECN3 site:

- **4 TEL62** readout boards (two used during last run)
- **Switch** HP2920 switch
- **Dual socket server** Intel Xeon E5-2620 @2.00 GHz CPUs
- **NANET-1 BOARD** (developed at INFN Roma APE Lab)
- **TTC HSMC** daughtercard (INFN Ferrara)
- **K20c** Kepler-class nVIDIA GPU.



GPU-based L0 trigger data flow



NaNet-1

GPU kernels consume CLOPs

Events from differents boards need to be merged before the ring-reconstruction stage.

KERNEL
Merging stage

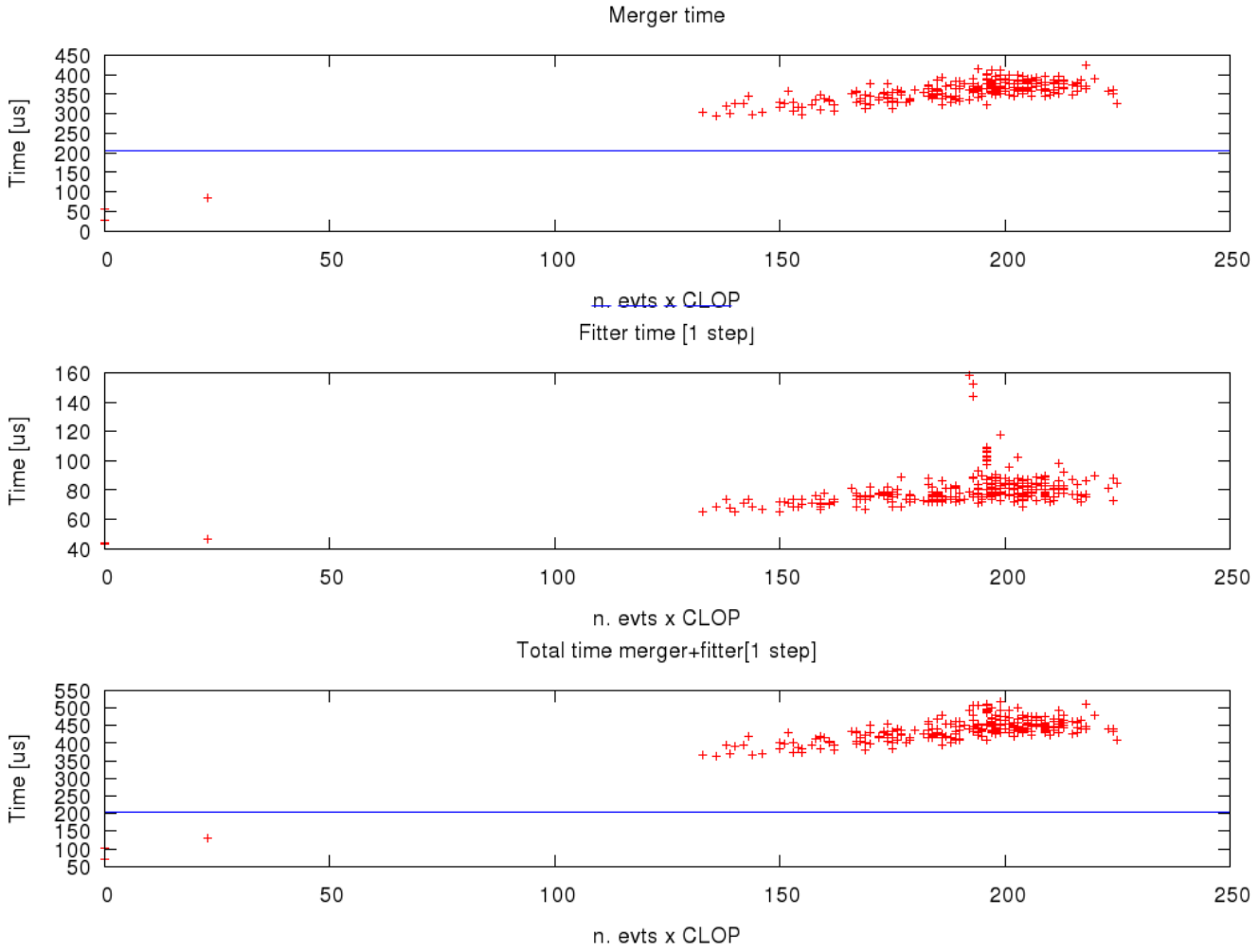
During the 2015 run it was used a Histogram based algorithm.
See *Stefano Di Lorenzo* talk!!

KERNEL
Ring-reconstruction

Inside NaNet-1 data coming from the 2 TEL62 boards are decompressed, rearranged ...

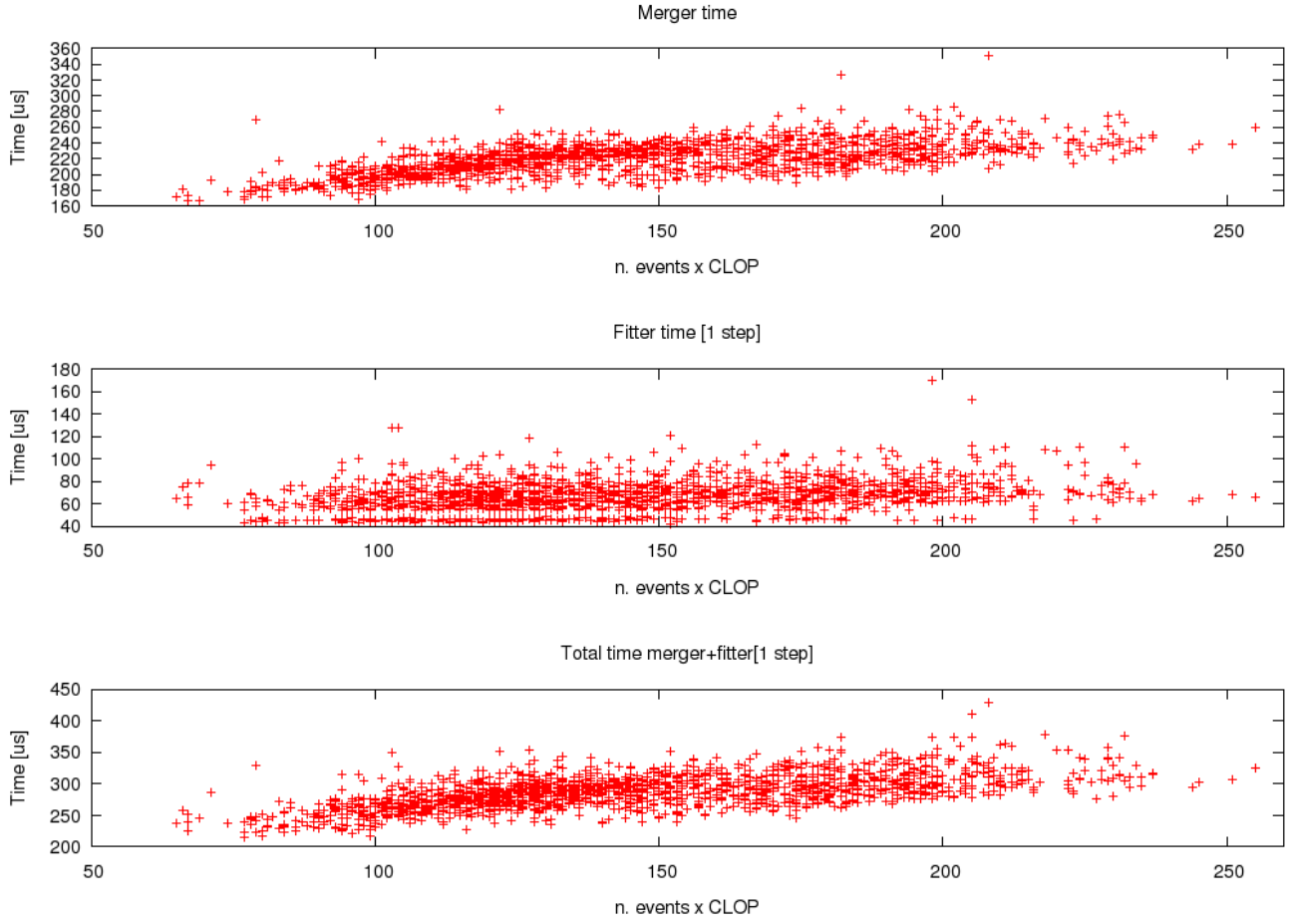
NA62 2015 Run: some results

Beam intensity: 2.4*10^6 pps [22/10/2015]



NA62 2015 Run: some results

Beam intensity: 33*10^6 pps [14/11/2015]

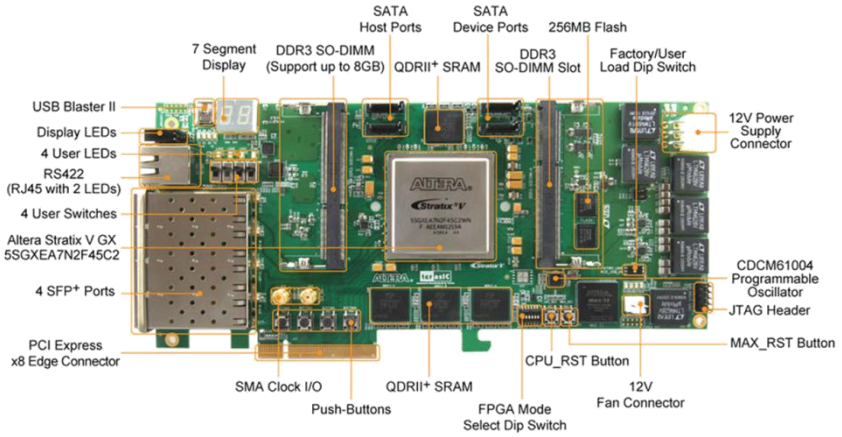


Looking forward to NA62 2016 Run

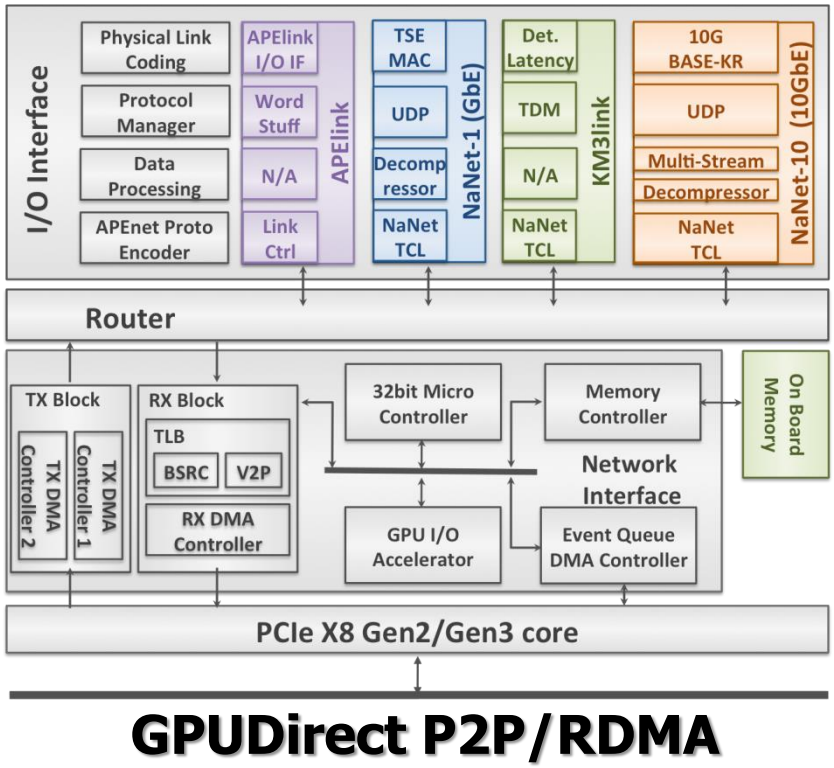
Looking forward to the 2016 Run: NaNet-10

To cope with the beam full intensity and to work with all of the 4 TEL62 boards...

- ALTERA Stratix V dev board
- PCIe x8 Gen3 (8 GB/s)
- 4 SFP+ ports (Link speed up to 10Gb/s)
- Implemented on Terasic DE5-NET board
- GPUDirect /RDMA capability
- UDP offloads supports

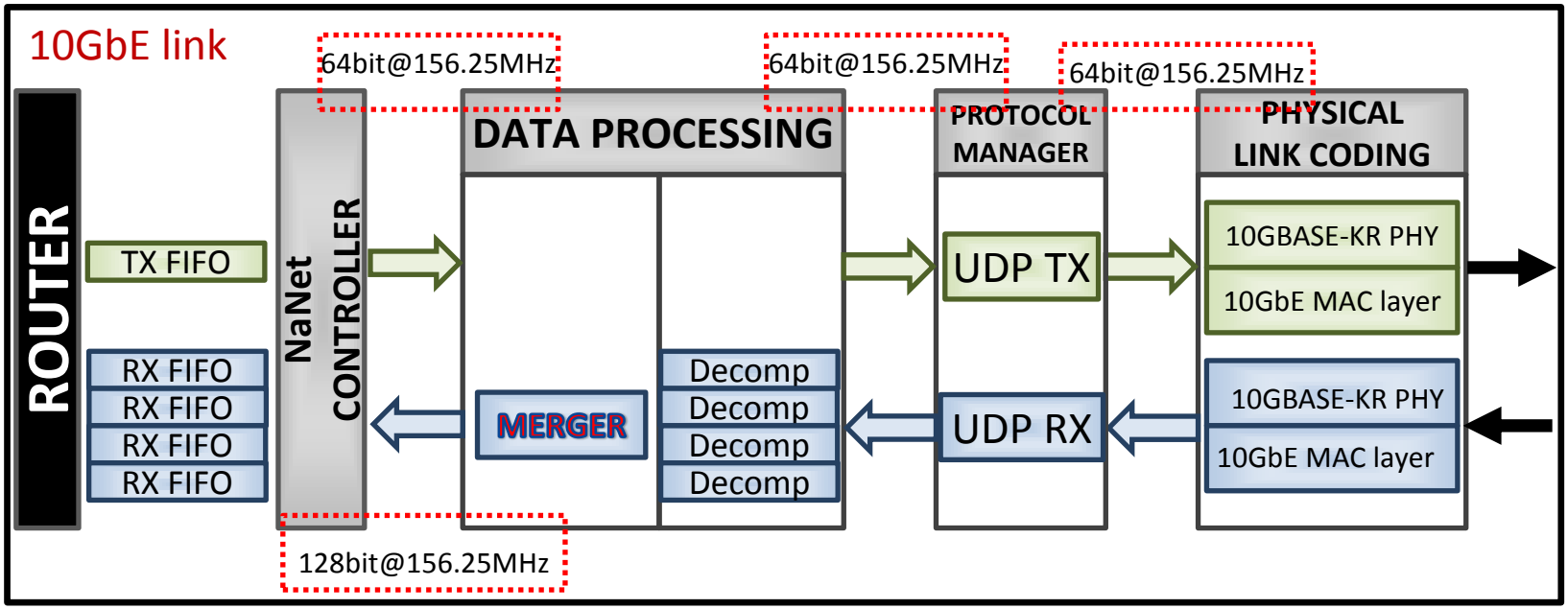


NaNet: Reconfigurable Design of Low-Latency NICs



- I/O Interface
  - Multiple physical link technologies.
  - Network protocols offloading.
  - Application-specific processing on data stream.
- Router
  - Dynamically interconnects I/O and NI ports.
- Network Interface
  - Manages packets TX/RX from and to CPU/GPU memory.
  - Zero-Copy RDMA.
  - GPU I/O accelerator.
  - TLB for Virtual to Physical mem map.
  - Microcontroller.
- PCIe X8 Gen2/3 Core

Looking forward to the 2016 Run: NaNet-10



Looking forward to the 2016 Run: EVENTS MERGER

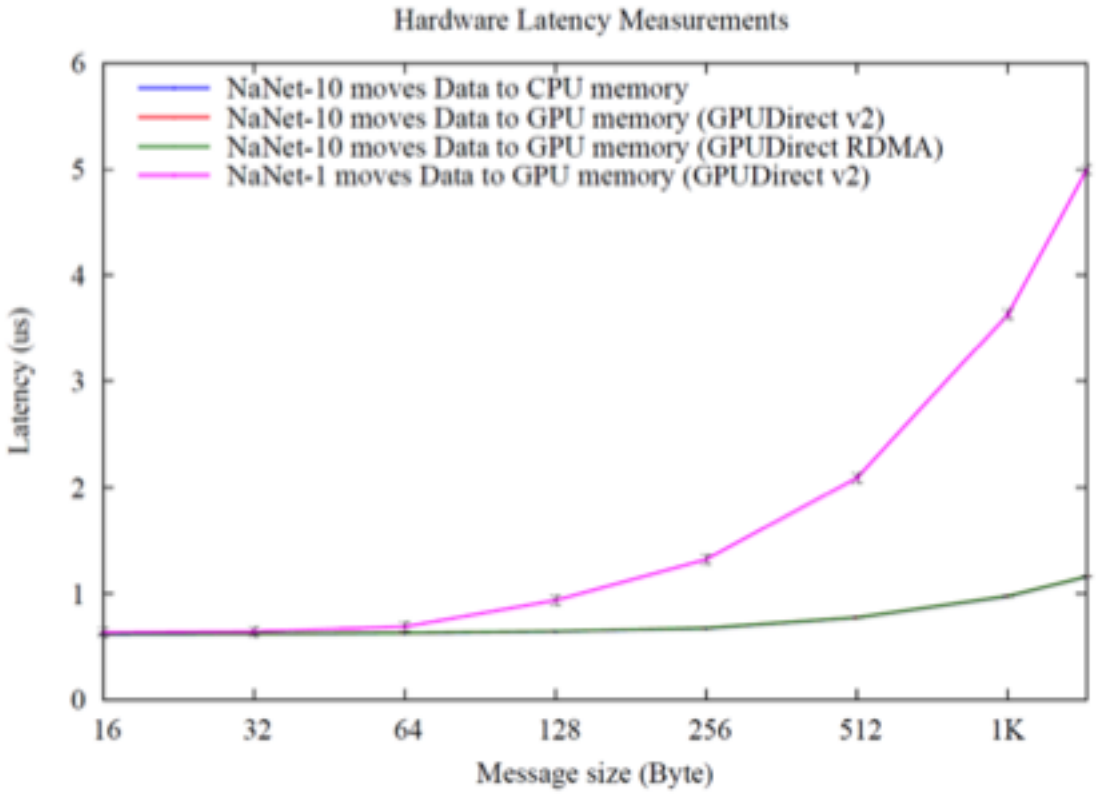
- **Merging stage will be performed in HW**
- **Events are arranged in CLOPs with a new format more suitable for GPU's threads memory access Multi Merged Event GPU Packet (M^2EGP).**

STR 3 MGP	STR 2 MGP	STR 1 MGP	STR 0 MGP	STR 3 HIT	STR 2 HIT	STR 1 HIT	STR 0 HIT	RESERVED	WINDOW	TOTAL HIT		TIMESTAMP			
STREAM 1; HIT 1		STREAM 1; HIT 0		STREAM 0; HIT 5		STREAM 0; HIT 4		STREAM 0; HIT 3		STREAM 0; HIT 2		STREAM 0; HIT 1		STREAM 0; HIT 0	
STREAM 2; HIT 0		STREAM 1; HIT 8		STREAM 1; HIT 7		STREAM 1; HIT 6		STREAM 1; HIT 5		STREAM 1; HIT 4		STREAM 1; HIT 3		STREAM 1; HIT 2	
STREAM 2; HIT 8		STREAM 2; HIT 7		STREAM 2; HIT 6		STREAM 2; HIT 5		STREAM 2; HIT 4		STREAM 2; HIT 3		STREAM 2; HIT 2		STREAM 2; HIT 1	
STREAM 3; HIT 4		STREAM 3; HIT 3		STREAM 3; HIT 2		STREAM 3; HIT 1		STREAM 3; HIT 0		STREAM 2; HIT 11		STREAM 2; HIT 10		STREAM 2; HIT 9	
PADDING										STREAM 3; HIT 7		STREAM 3; HIT 6		STREAM 3; HIT 5	
127…120	119…112	111…104	103…96	95…88	87…80	79…72	71…64	63…56	55…48	47…40	39…32	31…24	23…16	15…8	7…0

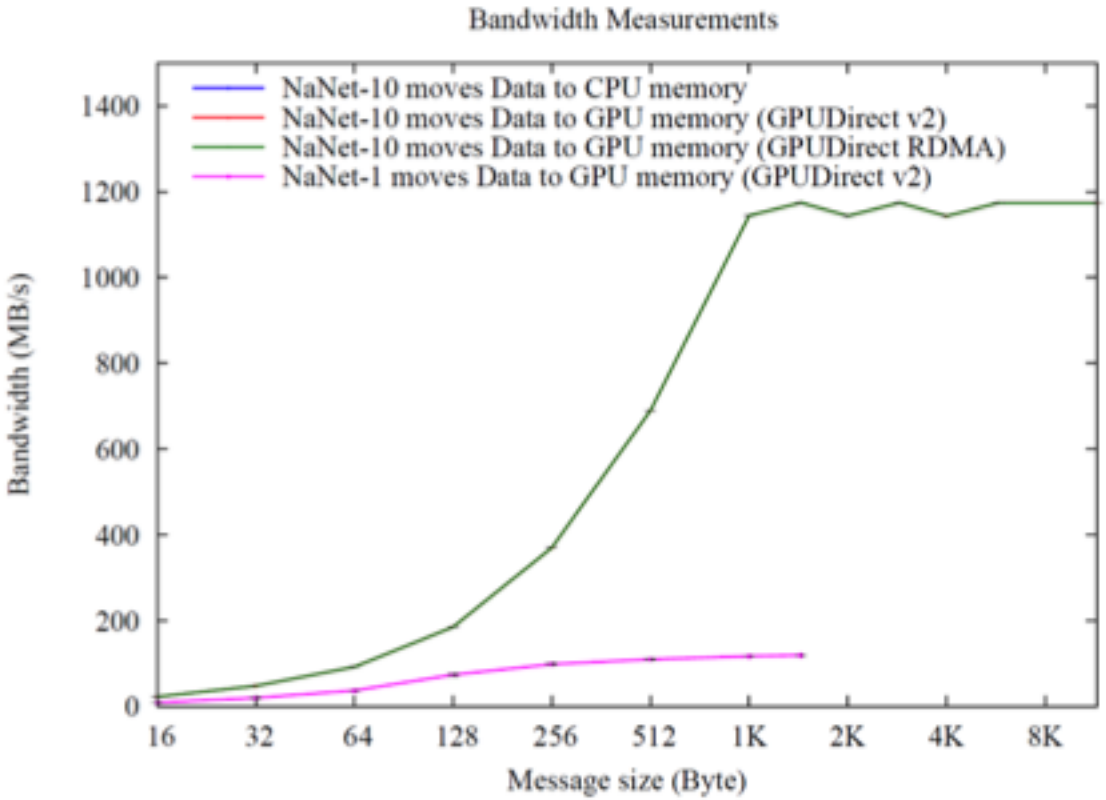
- **Problem: searching for events position inside a CLOP using 1 thread on GPU takes $>$ 100us for hundreds of events**
- **Solution: it must be parallelized. We can use all the threads looking for a known bytes pattern at the begin of every event: it takes $\sim$ 35us for 1000 events in a buffer**

[Thanks to Alessandro Lonardo for the advice]

Performances: NaNet-1 vs. NaNet-10 Latency



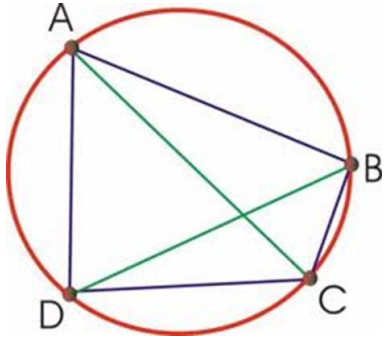
Performances: NaNet-1 vs. NaNet-10 Bandwidth

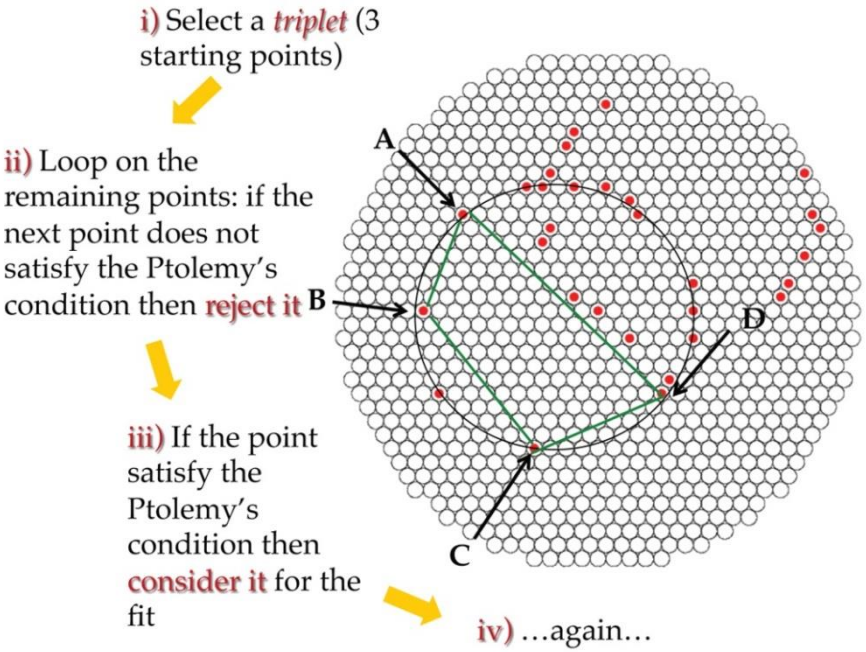


Looking forward to the 2016 Run: ALMAGEST new version

Based on Ptolemy's theorem:

"A quadrilater is cyclic (the vertex lie on a circle) if and only if is valid the relation:
AD*BC+AB*DC=AC*BD"





- **Several triplets run in parallel**
- **Several events at the same time**

In development a new kernel version for the 2016 RUN, leveraging the different merged data format and the recent NVIDIA GPU features: *warp vote functions, inter-threads communication*

Quick summary

- Installation of NaNet-10 at CERN *(first week of April)*
  with new system for receiving TTC signals (see *Ilaria Neri* talk!!*)*
- Test of the new setup
- Work in progress on the kernels side to:
  - adapt them to the new merged data format
  - test a new version of the Almagest based kernel leveraging on new CUDA instructions available for boards with Compute Capability > 3.0
- Start of 2016 Run on 25th of April

THANKS to:

**R. Ammendola [(b)], A. Biagioni[(a)], S. Chiozzi[(d)],
A. Cotta Ramusino[(d)], S. Di Lorenzo[(f,g)] , R. Fantechi[(f)],
M. Fiorini[(d,e)], O. Frezza[(a)], G. Lamanna[(c)],
F. Lo Cicero[(a)], A. Lonardo [(a)], M. Martinelli[(a)],
I. Neri[(d)], P.S. Paolucci[(a)], E. Pastorelli[(a)],
R. Piandani[(f)], F. Simula[(a)], M. Sozzi[(f,g)], P. Vicini[(a)].**

[(a)] INFN Sezione di Roma
[(b)] INFN Sezione di Roma Tor Vergata
[(c)] INFN - Laboratori Nazionali di Frascati
[(d)] INFN Sezione di Ferrara
[(e)] Università di Ferrara
[(f)] INFN Sezione di Pisa
[(g)] Università di Pisa