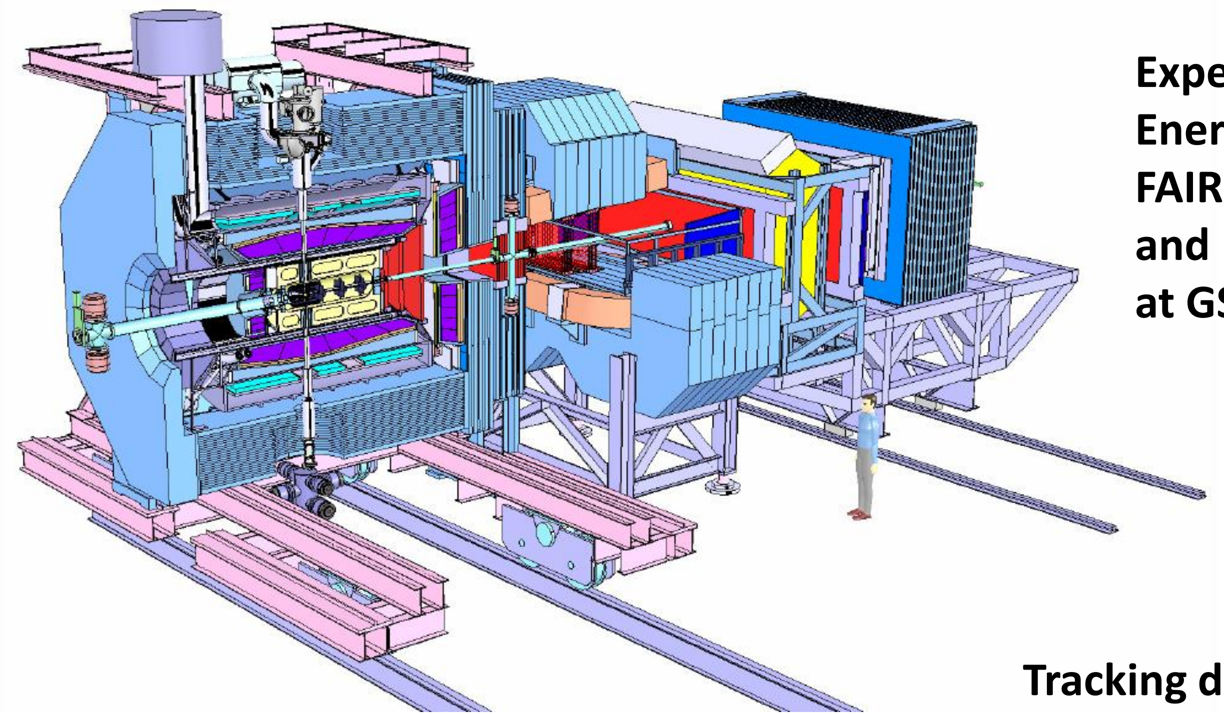
Modeling event building architecture for the triggerless data acquisition system for PANDA experiment at the HESR facility at FAIR/GSI

Krzysztof Korcyl
Institue of Nuclear Physics Polish Academy of Sciences
&
Cracow University of Technology

Agenda

- PANDA experiment – detectors and requirements for DAQ system
- the push-only architecture
- Compute Node in ATCA standard
- data flow in the architecture
- short introduction on discrete event modelling
- modeling results
  - latency
  - queues dynamics
  - load monitoring
- summary

PANDA detectors



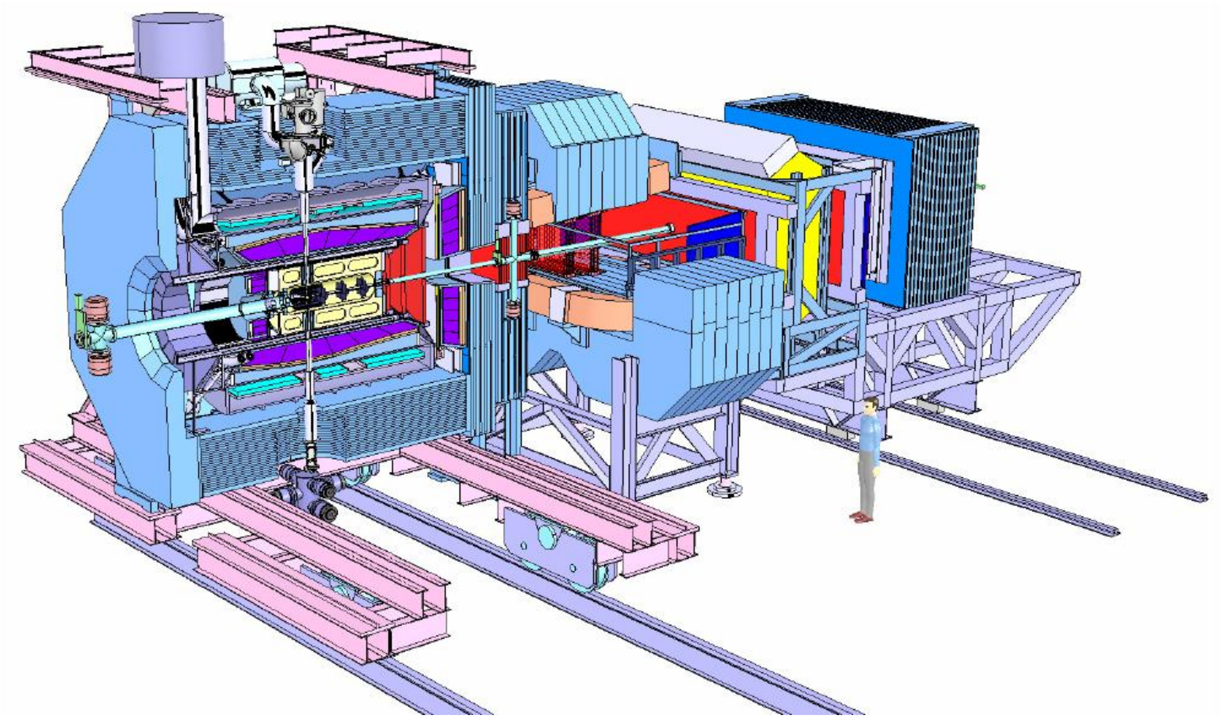
Experiment at HESR (High Energy Storage Ring) in FAIR (Facility for Antiproton and Ion Research) complex at GSI, Darmstadt.

Tracking detectors:
- Micro Vertex Detector
- Central Tracker
- Gas Electron Multiplier Stations
- Forward Tracker

Calorimetry:
- Electromagnetic calorimeter

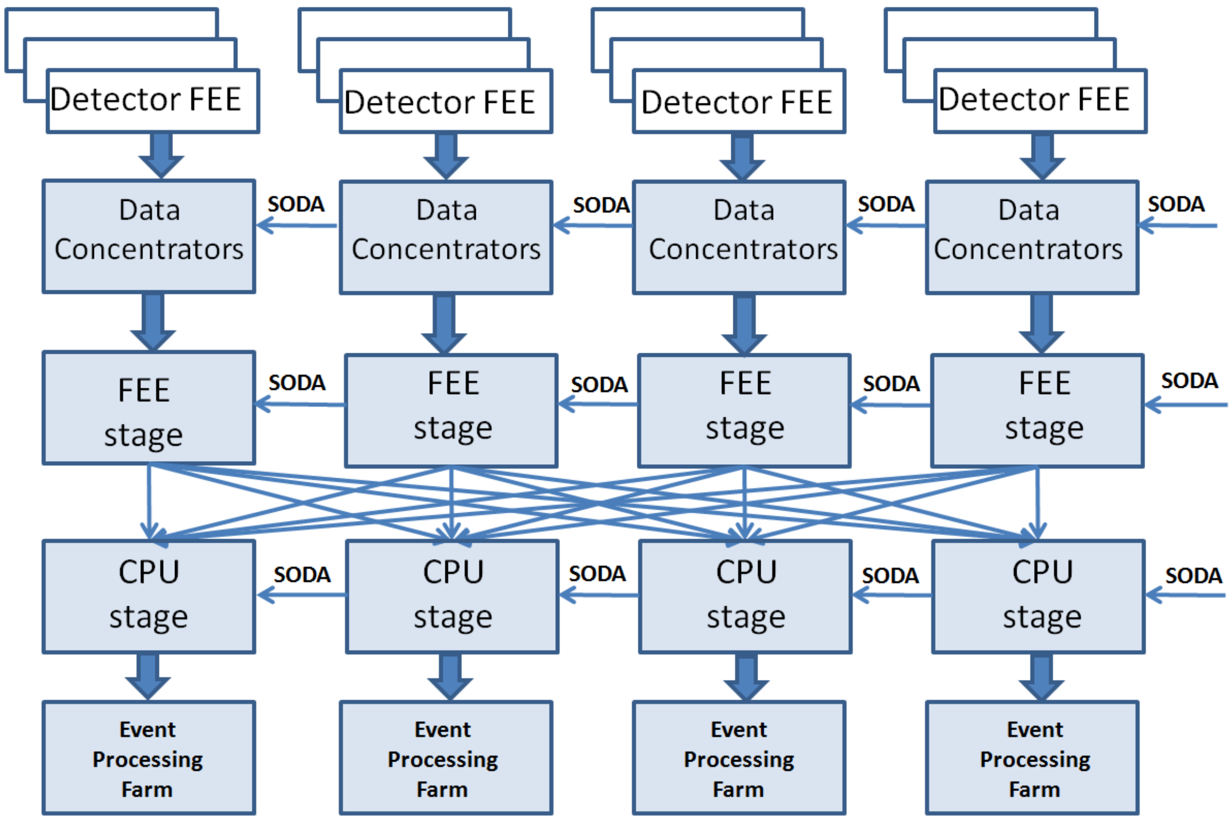
Particles identification:
- DIRC (Detection of Internally Reflected Cherenkov)
- Time of Flight System
- Muon Detection System
- Ring Imaging Cherenkov Detector

PANDA DAQ requirements



- interaction rate: up to 20 MHz (luminosity $2 * 10^{32}$ $cm^{-2}s^{-1}$)
- typical event size : ~4 kB
- expected throughput: 80 GB/s (100 GB/s)
- rich physics program requires a high flexibility in event selection
- front end electronics working in continuous sampling mode
- lack of hardware trigger signal

The push-only architecture

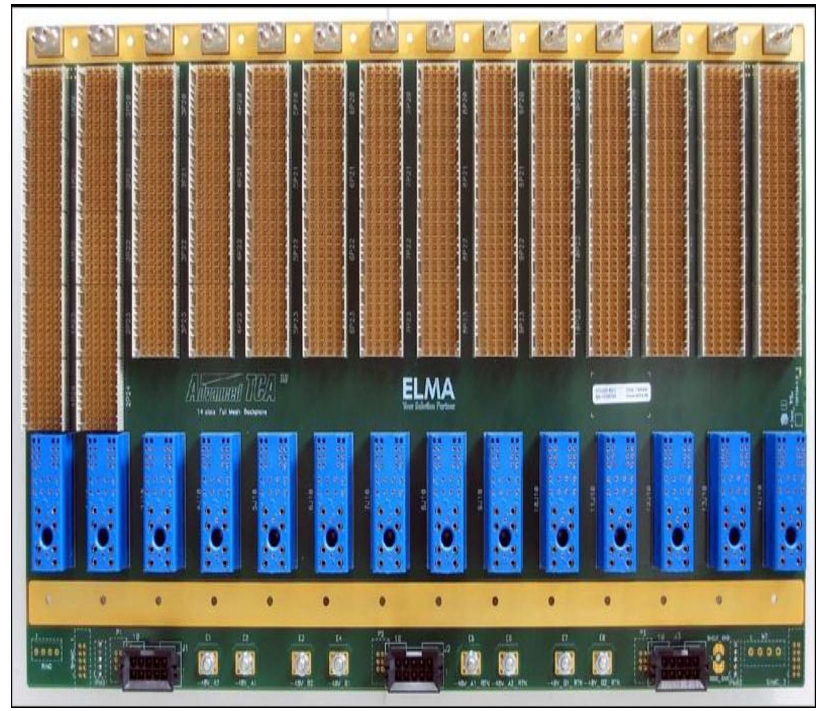


SODA

Time Distribution System

Passive poin-to-multipoint bidirectional fiber network providing time reference with precission better than 20 ps and performs synchronization of data taking

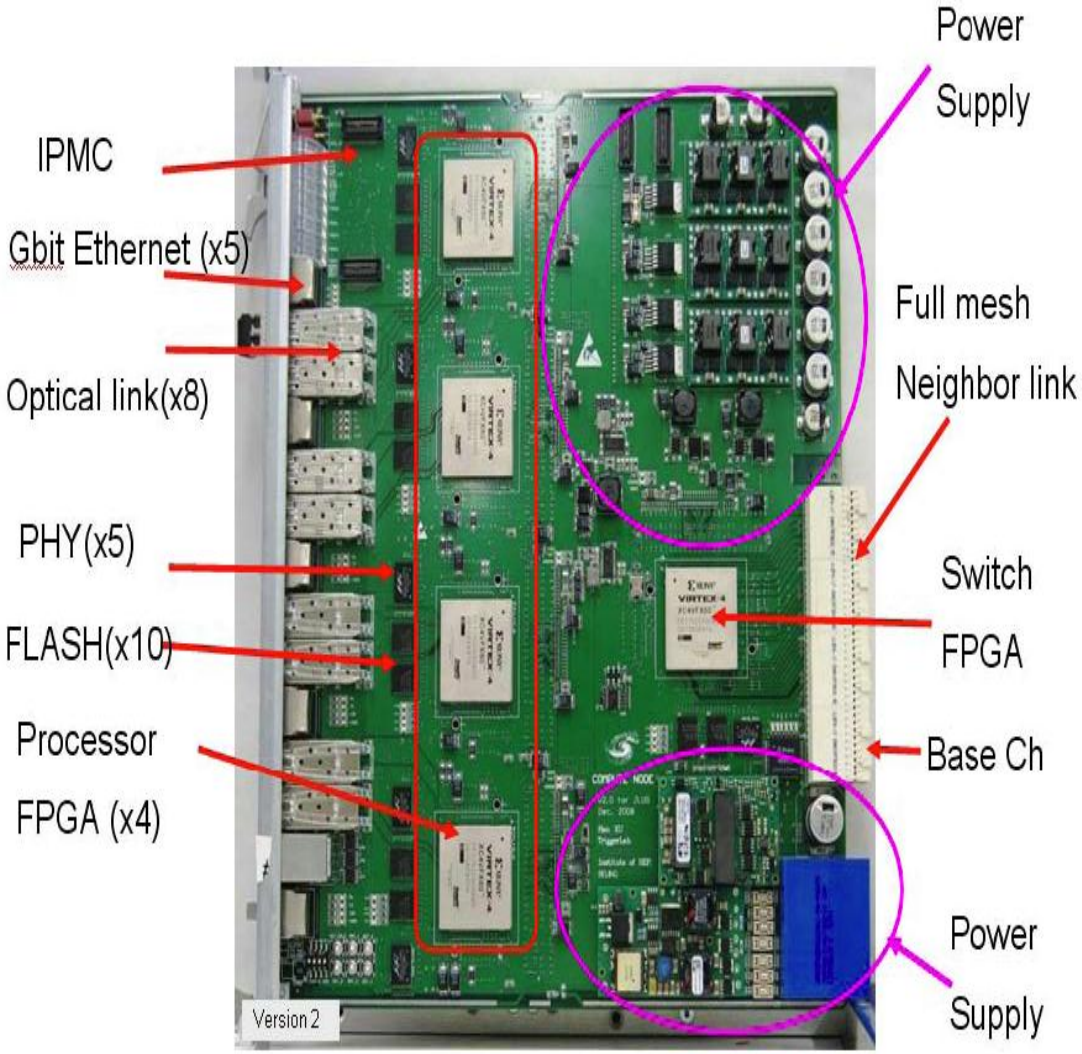
ATCA crate and backplane



ATCA – Advanced Telecommunications Computing Architecture

Backplane: one of possible configuration is full mesh (connects each pair of modules with dedicated point-to-point bidirectional llink

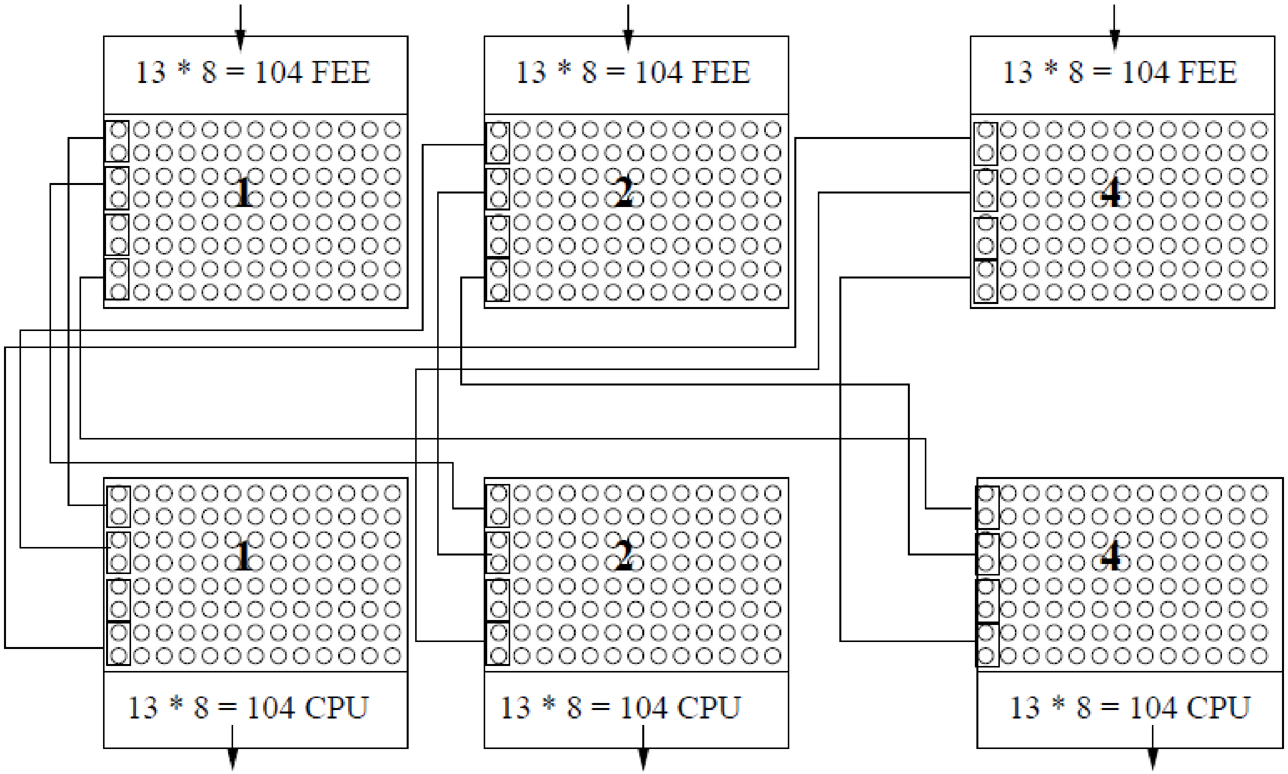
Compute Node

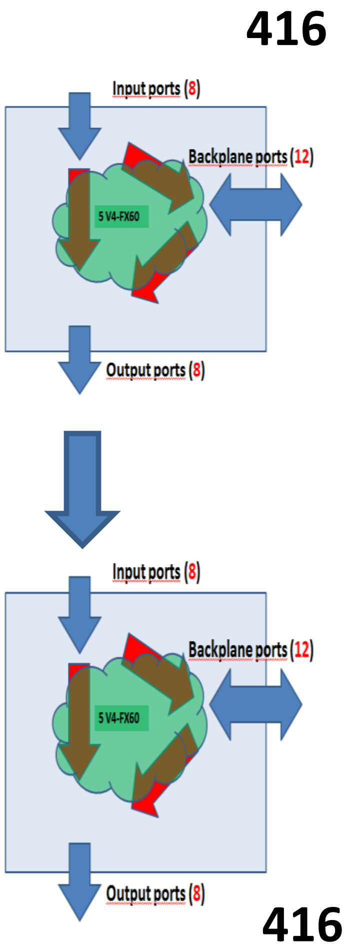


Each board is equipped with 5 Virtex4 FX60 FPGAs.

High bandwidth connectivity is provided by 8 Gbit optical links connected to RocketIO ports (6.5 Gb/s).
In addition the board is equipped with five Gbit Ethernet links

Inter-crate wiring

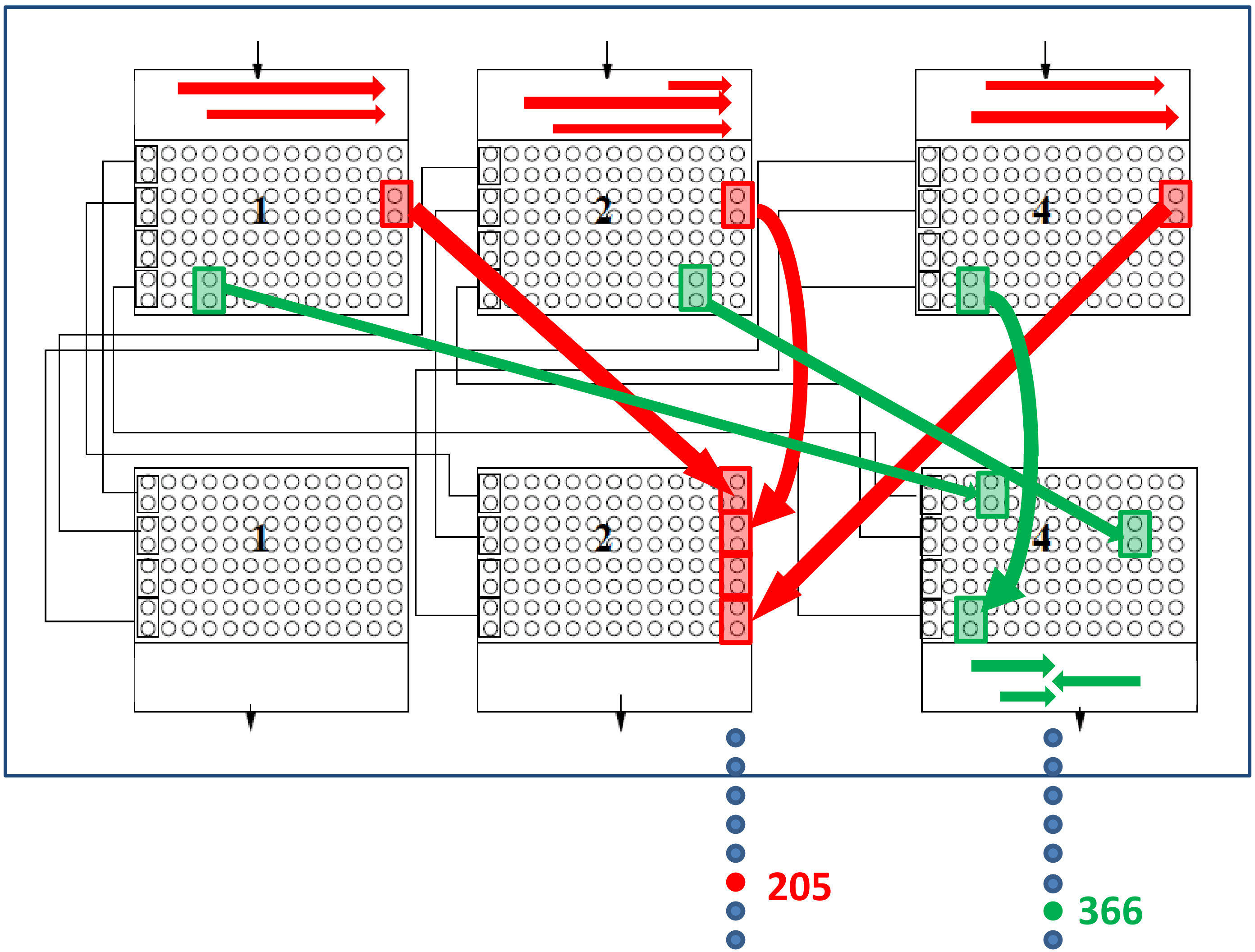




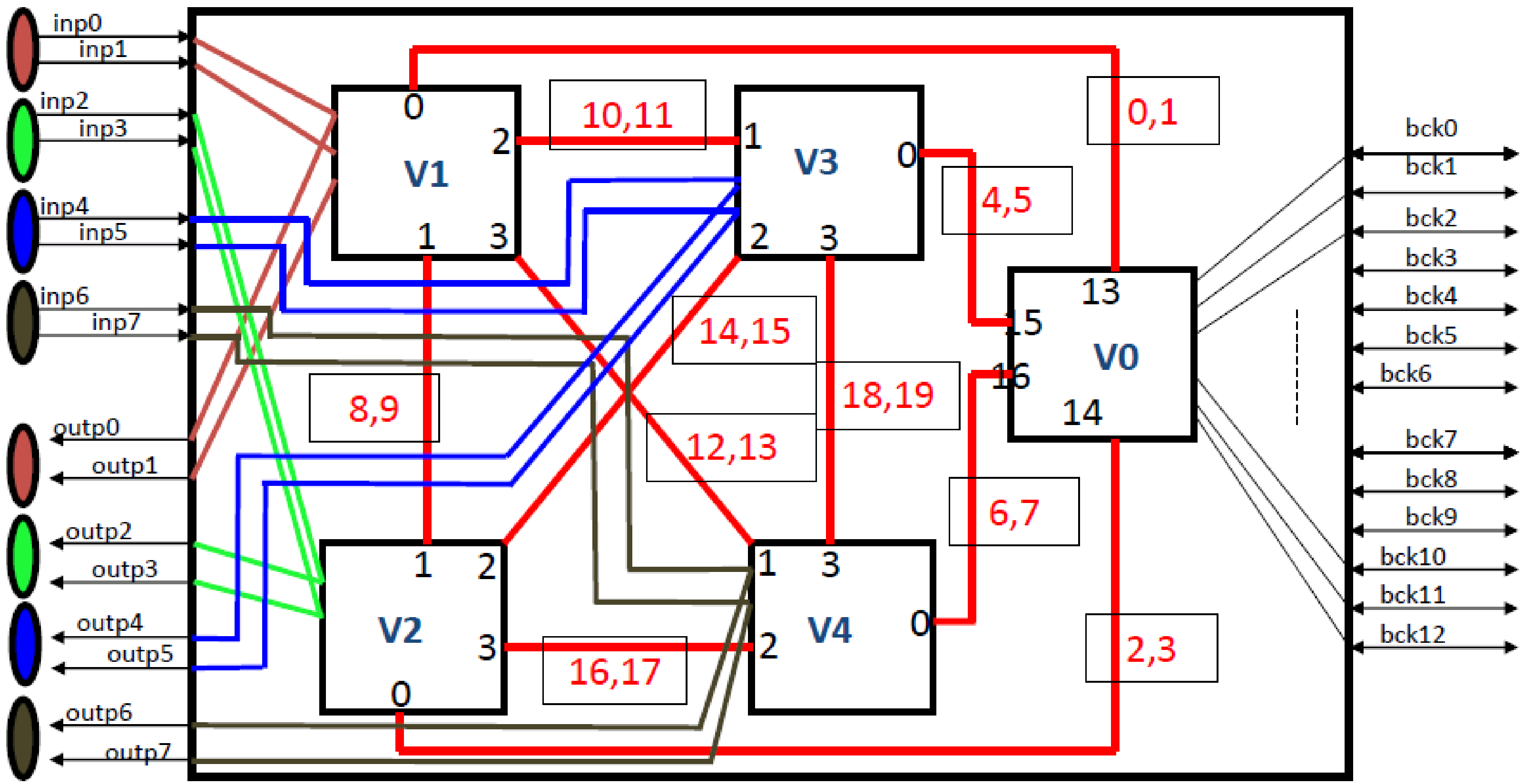
Module in slot N at the **FEE** level connects, with 2 links trunks, to modules at slots N at the **CPU** level.

The odd events packets at the FEE level are first routed via the backplane and then outbound to the CPU level via a proper trunk. The even events packets outbound to the CPU level first and then use backplane to change the slot.

Inter-crate routing animation



Onboard Virtex connections

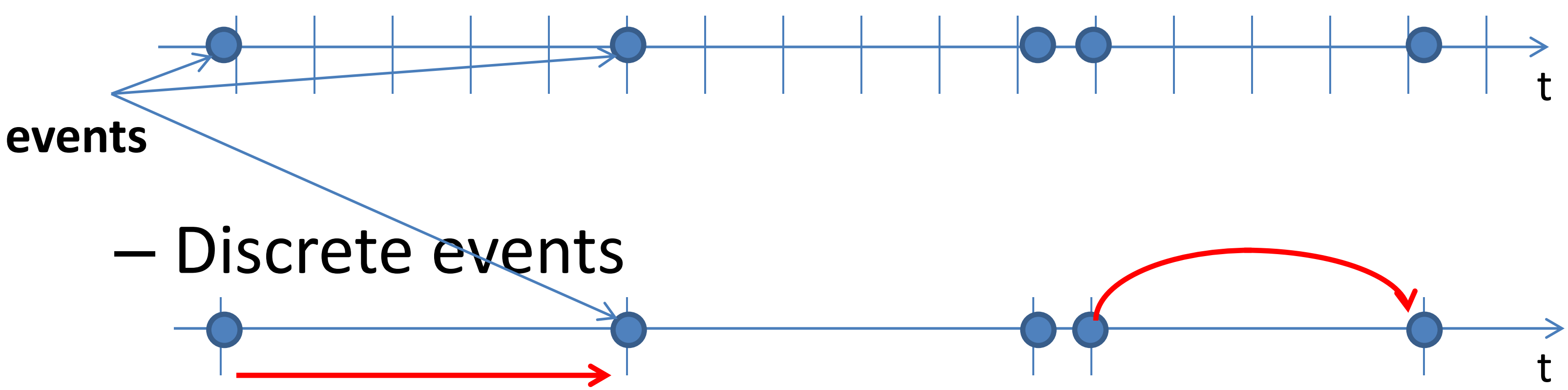


Virtex0 – handles all communications to/from the backplane

Virtex1-4 - manage 2 input and 2 output ports at the front panel

Discrete event modeling

- Model – computer program simulating system dynamics in time (support from SystemC library)
  - Fixed time step

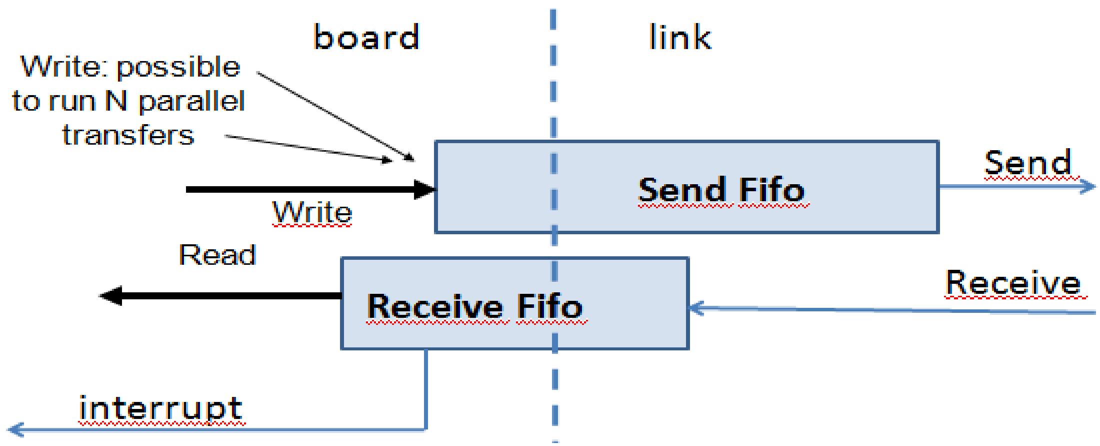


  - Discrete events

State of the system remains constant between events

Processing system in a state may lead to scheduling a new event in the future

Parameterization of ports



SendFifo – occupation can grow if multiple writes are allowed OR the link speed is smaller than the speed of write OR the recent packets are smaller than the former ones.

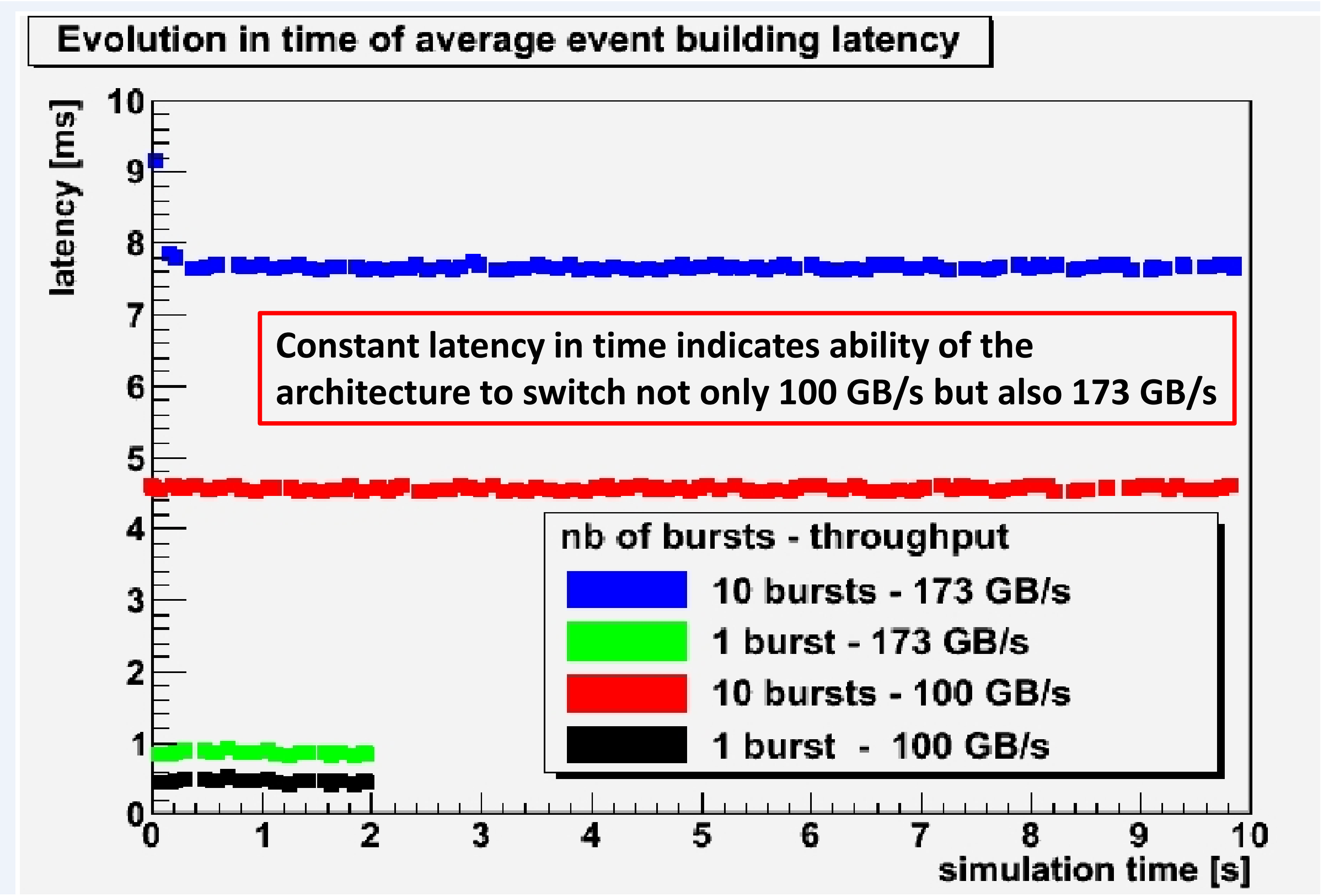
ReceiveFifo – occupation can grow if the queue head can not be transferred due to destination being busy with another transfer OR the recent packets are smaller than the former ones.

The transfer speed is a parameter – during the simulations it was set to 6.5 Gb/s (RocketIO)

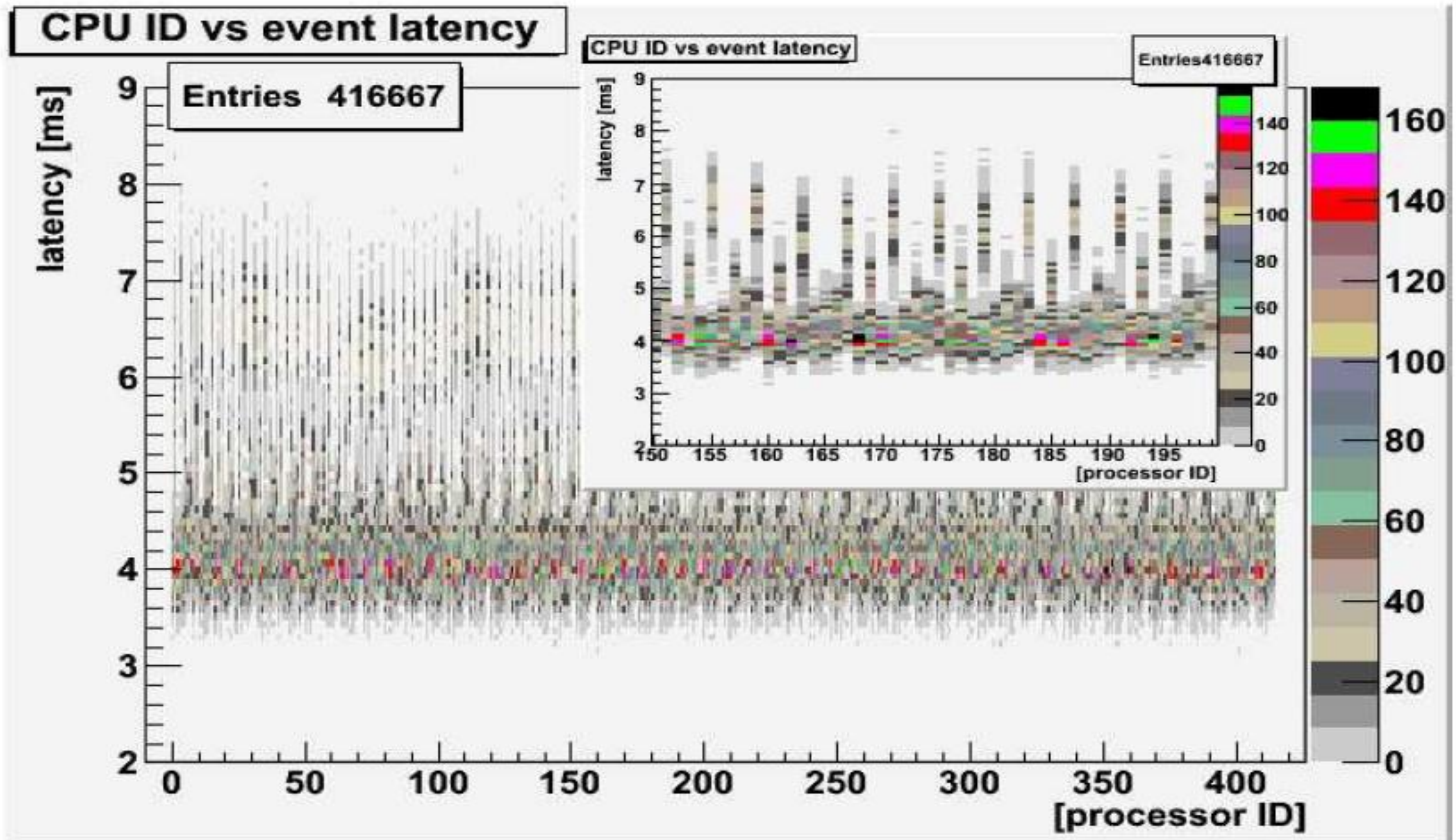
Models of source and sink of data

- Data source – Data Concentrator:
  - generates data packet with a size proportional to the sum of number of inter-interactions calculated from Poisson distribution with average of 20 MHz
    - Burst: 2 μs of interactions + 400 ns of silence gap
    - Super-burst: 10 bursts
  - Simulates the 8b/10b conversion, tags packets with destination CPU number and pushes into the architecture.
- Data sink – event building CPU:
  - Simulates event building of 416 fragments with the same tag
    - size of burst: ~300 kB
    - size of super-burst: ~3 MB

Event building latency



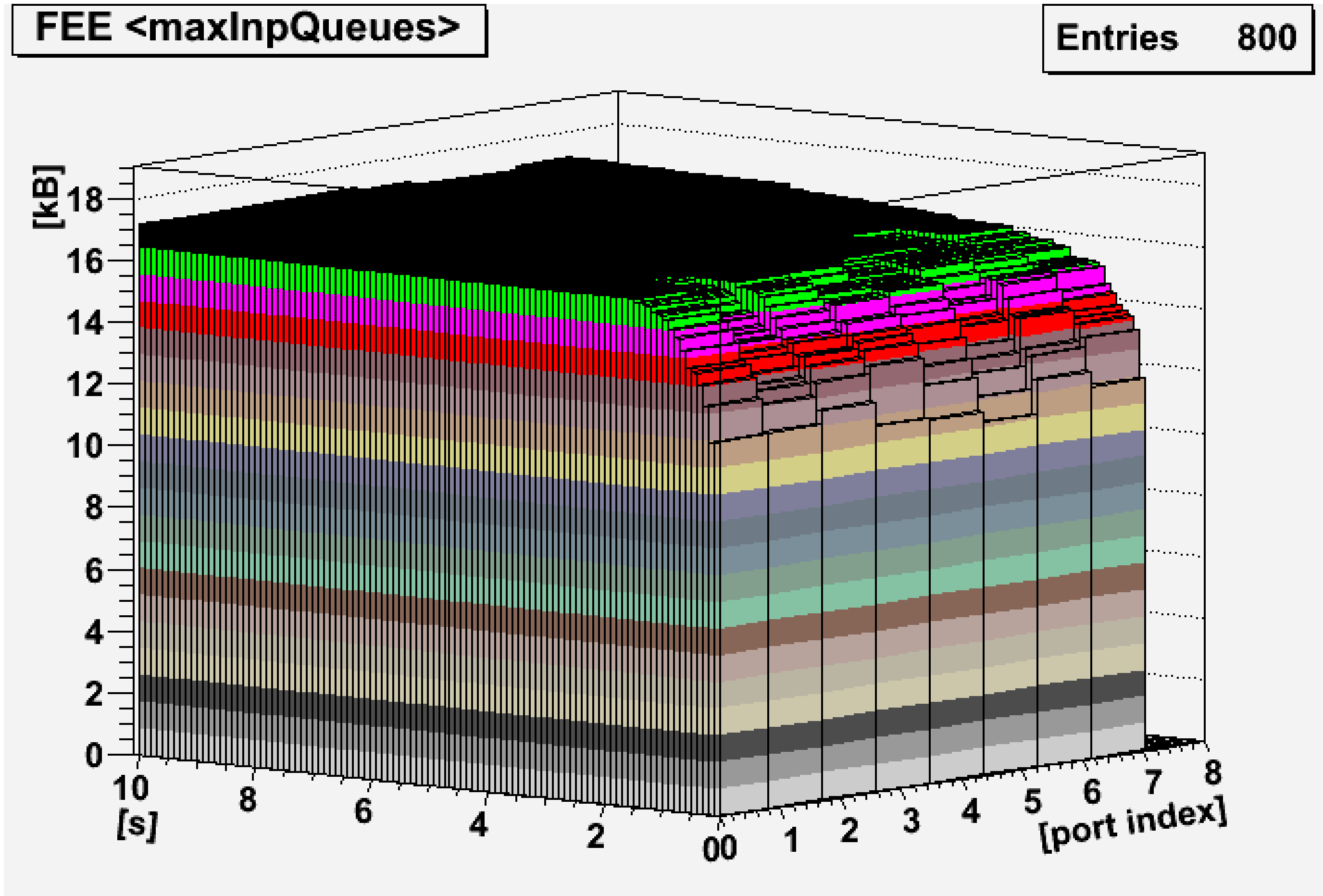
Load distribution between CPUs



The CPU for next event is calculated with the formula:

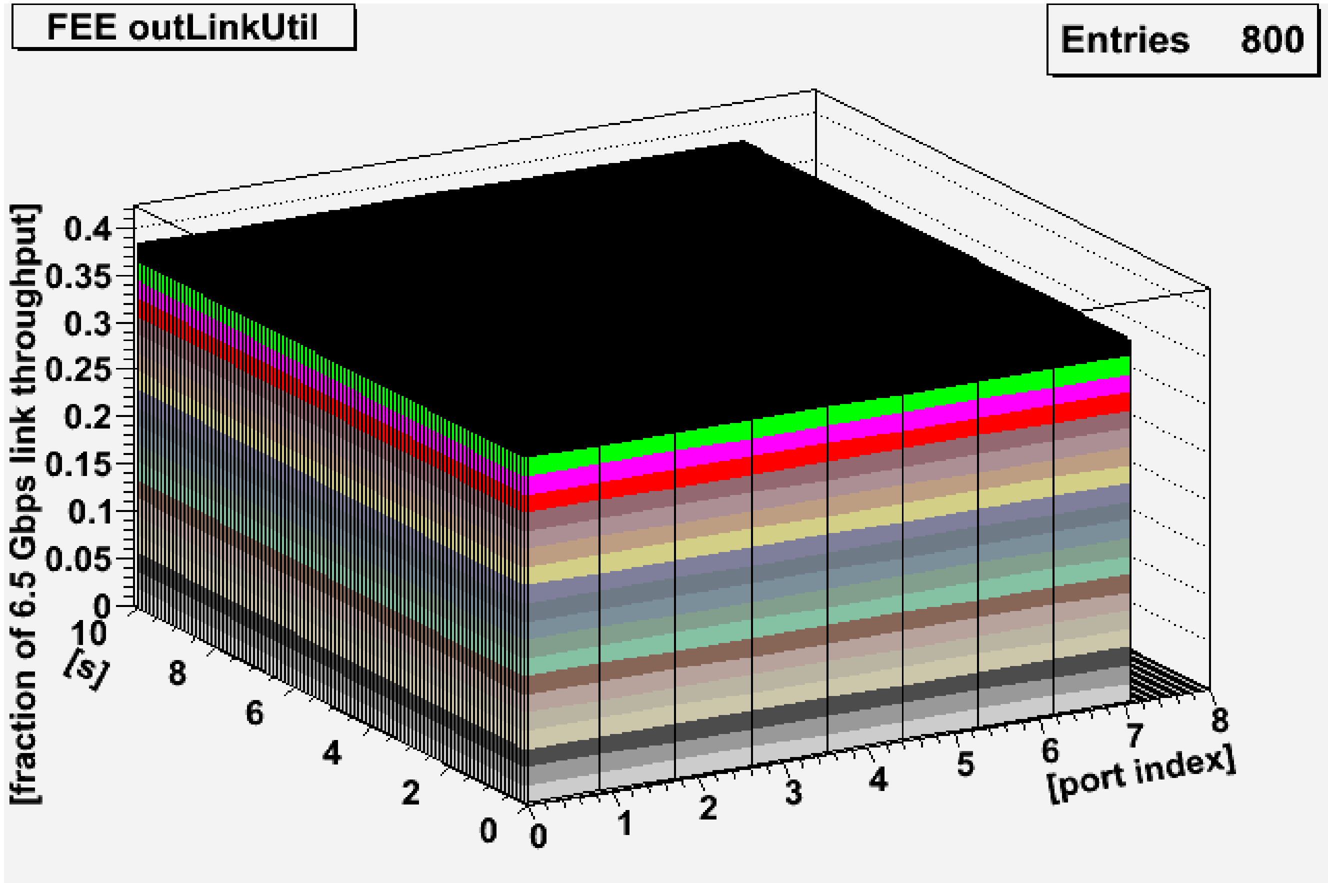
$$N_{t+1} = \mathrm{mod}\ (N_t + 79)\ ,\ 416)$$

Monitoring queues' evolution



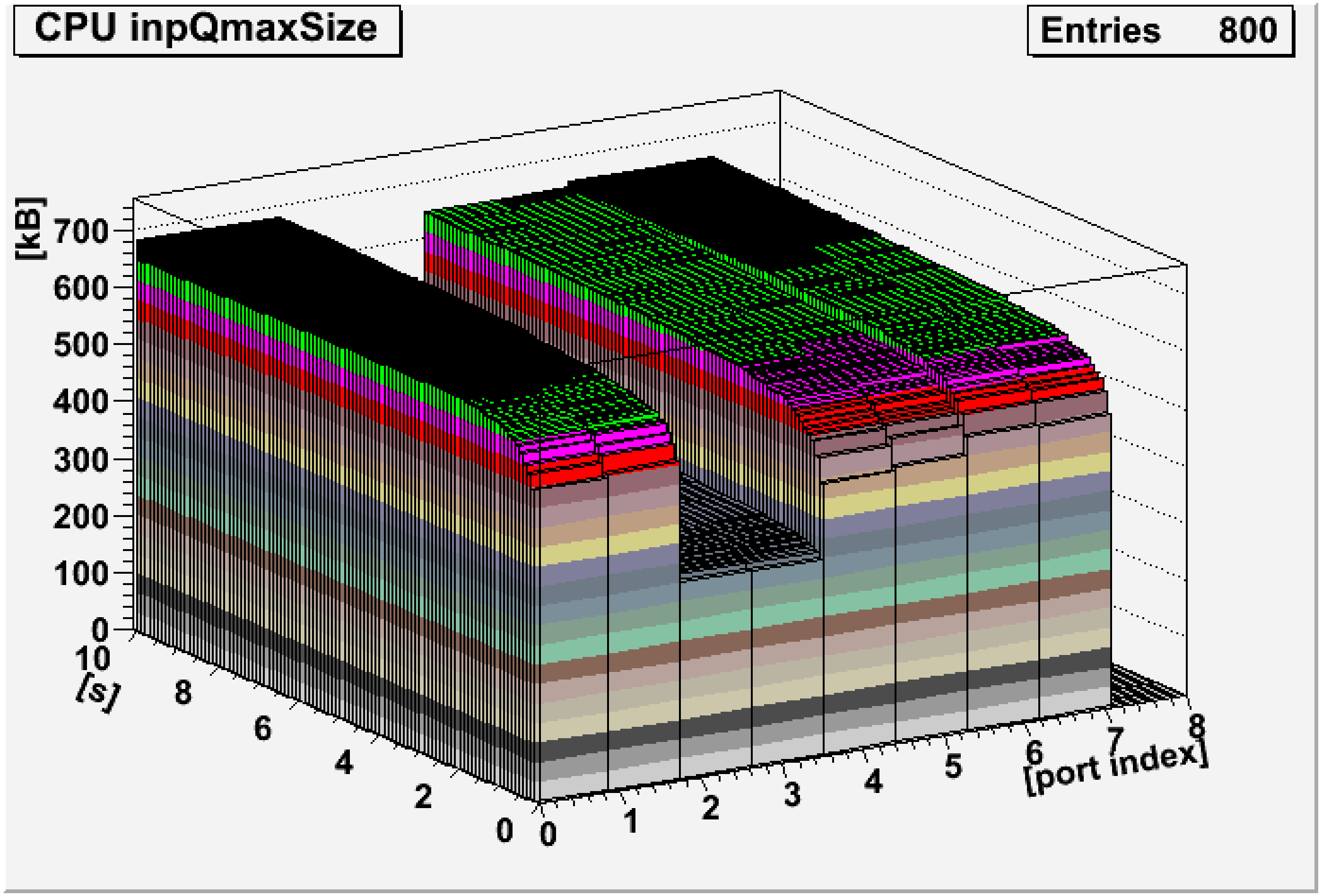
Averaged maximal queue length in input ports at the FEE level. The averaging was over the ports with the same index in all FEE modules.

Monitoring links' load



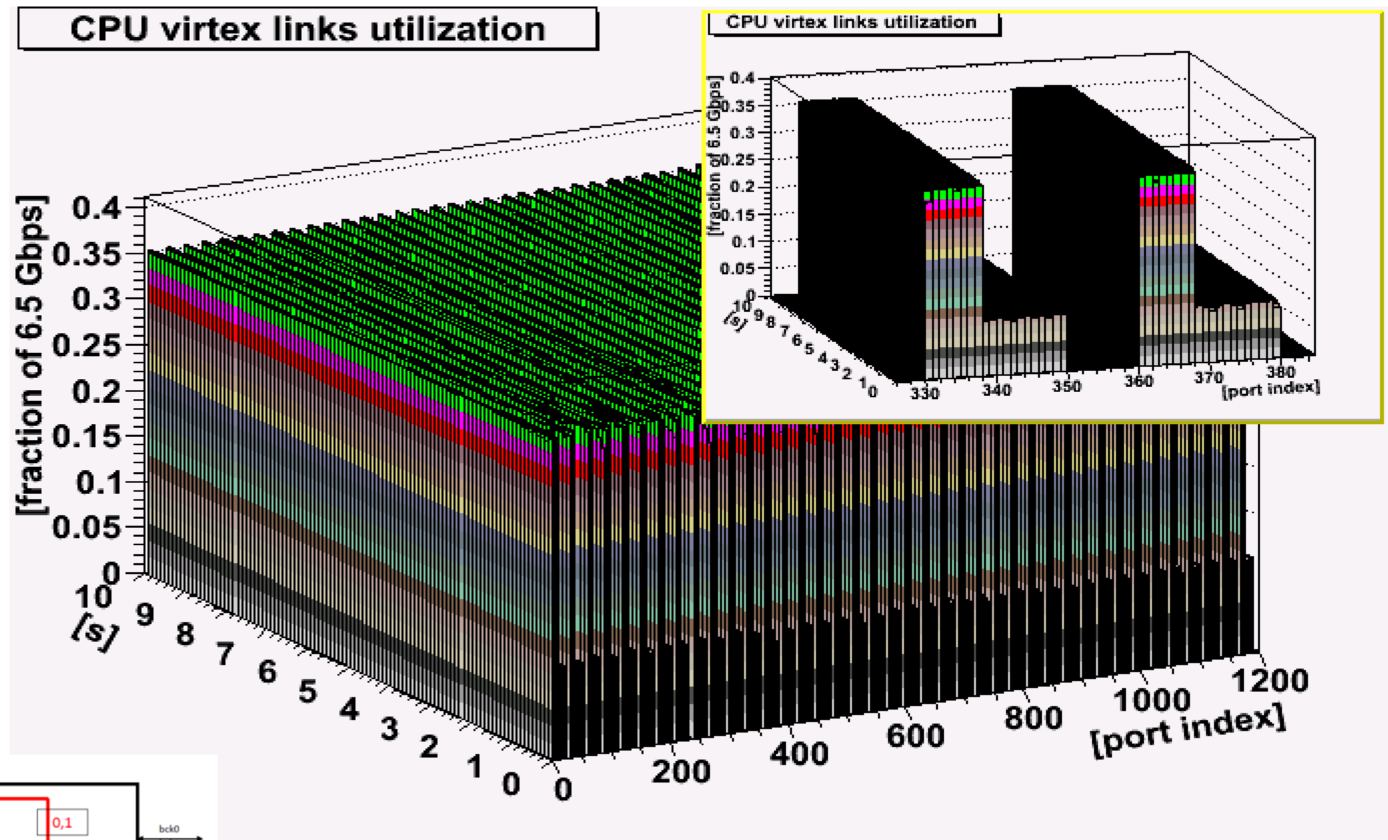
Load on fiber links connecting output ports from the FEE level with input ports at the CPU level. Homogenious load indicates proper routing scheme – also for trunking.

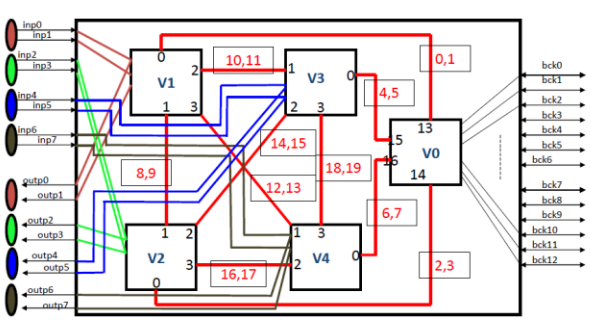
Monitoring queues



Average of maximal length of input queues at the CPU level. The average was made with ports with the same index on all modules.

Monitoring Virtex link's occupation





At the CPU level, the packets heading for odd-numbered CPUs go via the backplane to the slot with destination CPU.

Summary

- We propose the event building architecture for the triggerless DAQ system for the PANDA experiment.
- The architecture uses Compute Node modules in ATCA standard.
- We built simplified models of the components using SystemC library and run the full scale simulations to demonstrate required performance and to analyse dynamics of the system.
- The push-only mode offers 100 GB/s throughput which allows to perform burst/super-burst building and to run selection algorithms on fully assembled data.
  - with the input links loaded up to 70% of their nominal capacity the architecture can handle 173 GB/s