



Employing HPC for Heterogeneous HEP Data Processing

AUGUST 2018

AUTHOR:

Marco
Barbone

CERN IT-DI-OPL
DEEP-EST

SUPERVISORS:

Viktor Khristenko
Felice Pantaleo
Maria Girone





Project Specification

project specification





Abstract

THIS IS MY ABSTRACT





Contents

Contents	iv
List of Figures	v
List of Tables	vi
Listings	vii
1 From physics to... Physics	1
2 Second chapter	3
3 Conclusions	4
Bibliography	5
Appendix	6
A My First Appendix	6





List of Figures

1.1 Cern data flow from collisions to analysis	2
--	---





List of Tables





Listings



1. From physics to... Physics

Modern high energy physics (HEP) requires analysis. To make it possible several large scale objects are needed: from accelerators, detectors, data centers to the thousands of people involved to design and run the infrastructure. Here at CERN everything starts with physics by colliding particles and everything ends with the physics performed by analyzing the data acquired. But, to produce the necessary data, in the right amount, a long and complicated process, showed in figure 1.1, is involved. It is worth to briefly describe it to understand the reason of our work and where is it place inside it.

Data generation The process starts with **particle collisions** this lead to particle interactions that create some **intermediate product**. This intermediate product can not be directly observed, bit it decays in **particles** that go through the detectors. The intermediate product is divided in **X** and **Y**, X is what physicists are trying to study while Y is something that produces similar particles as Y but is not what they are looking for.

Data aquisition Detectors are split in two levels: **physics detector level** and **electronics level**. The physics level takes in input sensor readouts and produces an analog signal. This analog signal is the digitalized by the electronics level. This level also decides if the event is interesting or not, based on this the event is discarded or recorded. The event that survive this ends up dumped into **disks** and send to the **High level trigger**.

It is important to notice that these decision are taken in real time, to meet the throughput and timing requirements everything is built with **asics** and **fpgas** boards.

Data processing Data processing is performed both online and offline. The first step, called **local reconstruction**, in the offline data processing is to transform charges in physical quantities such as **energy**, **time**, and **position**. This operation is done channel by channel independently form each other. Then, information coming from different channel of the same detector are combined by the **cross channel clustering** into **local detector clusters**. The final step needed to obtain **particle objects** is the **cross detector clustering** in which data coming from different detector is combined. The particle object forms the dataset used by theoretical physicists to do physics.

High level trigger (HLT) The **HLT** belongs both to data acquisition and data processing, since it would not be correct to include it in one of this categories, it deserves a separate discussion. The HLT main purpose is to select interesting events before writing them to disk. In order to accomplish this, it runs the same code used for the offline processing, but **online** and with less features because of strict time constraints.

The main difference between **L1** and HLT is that L1 is implemented in hardware with no host while HLT is software based, it runs on hosts with accelerators used to speed-up the processing.



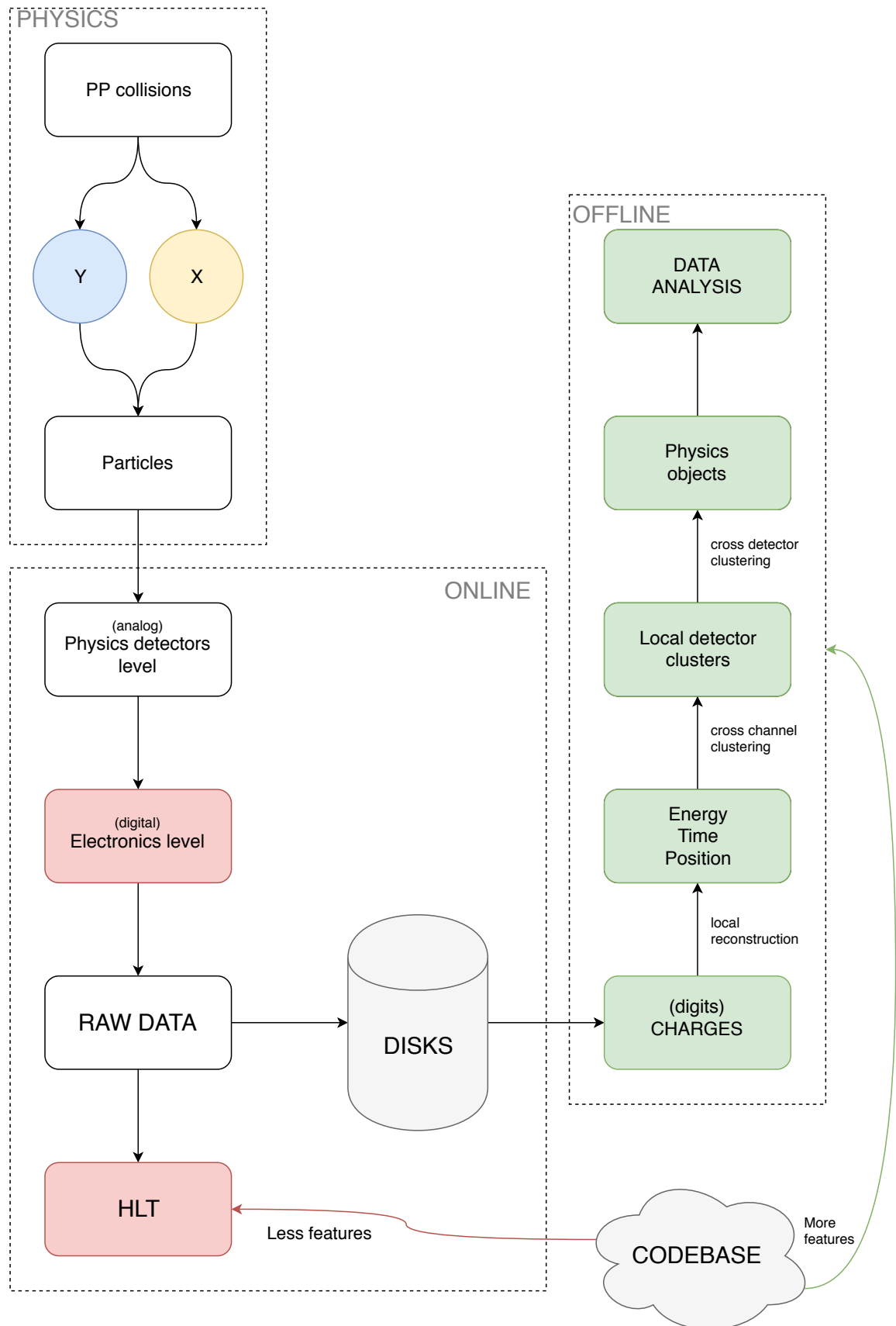


Figure 1.1: Cern data flow from collisions to analysis



2. Second chapter





3. Conclusions

Write your conclusions here.





Bibliography





A. My First Appendix

In this file (appendices/main.tex) you can add appendix chapters, just as you did in the thesis.tex file for the 'normal' chapters. You can also choose to include everything in this single file, whatever you prefer.

