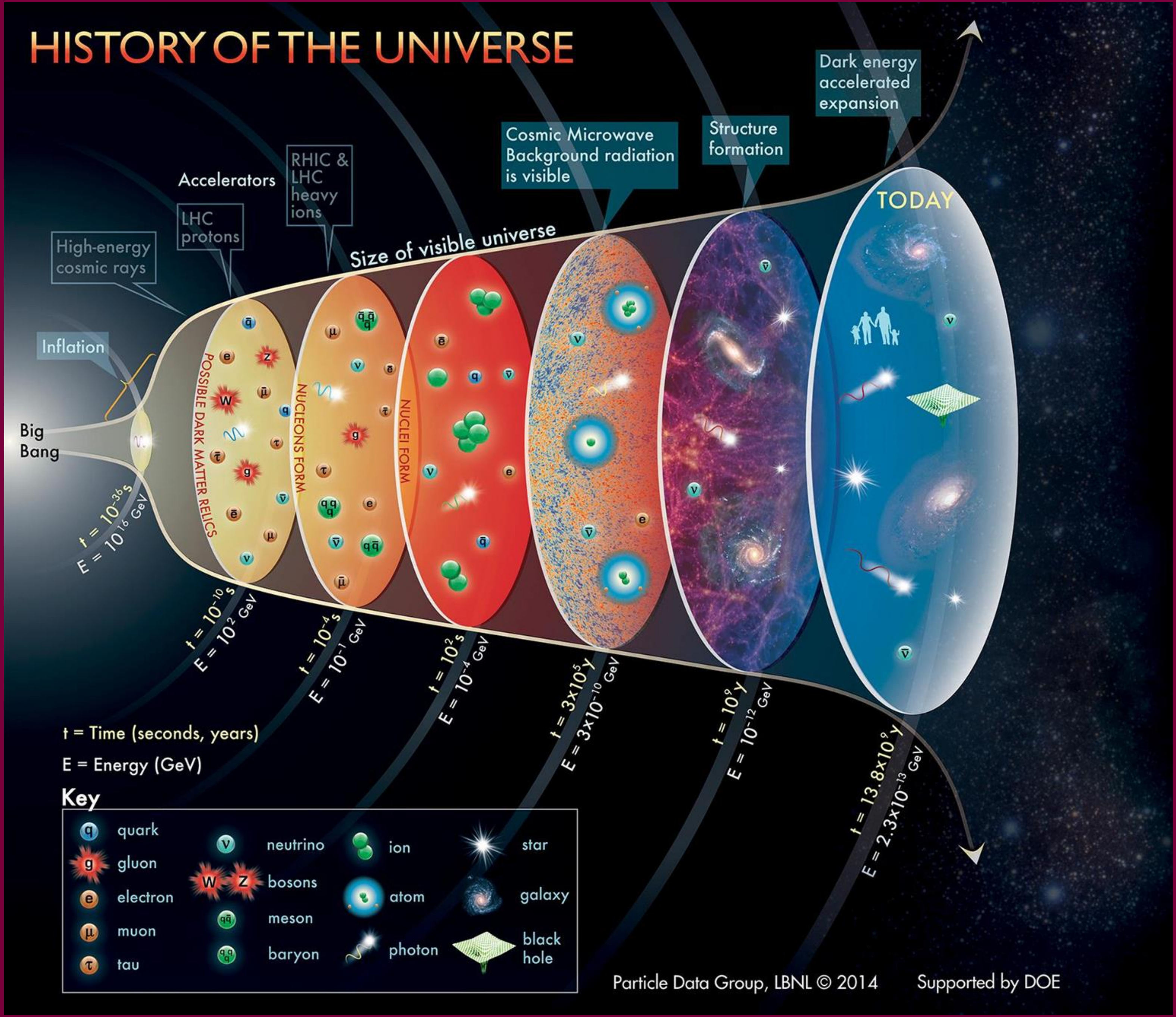
Vilniaus universitetas

Accelerators, detectors, and computing

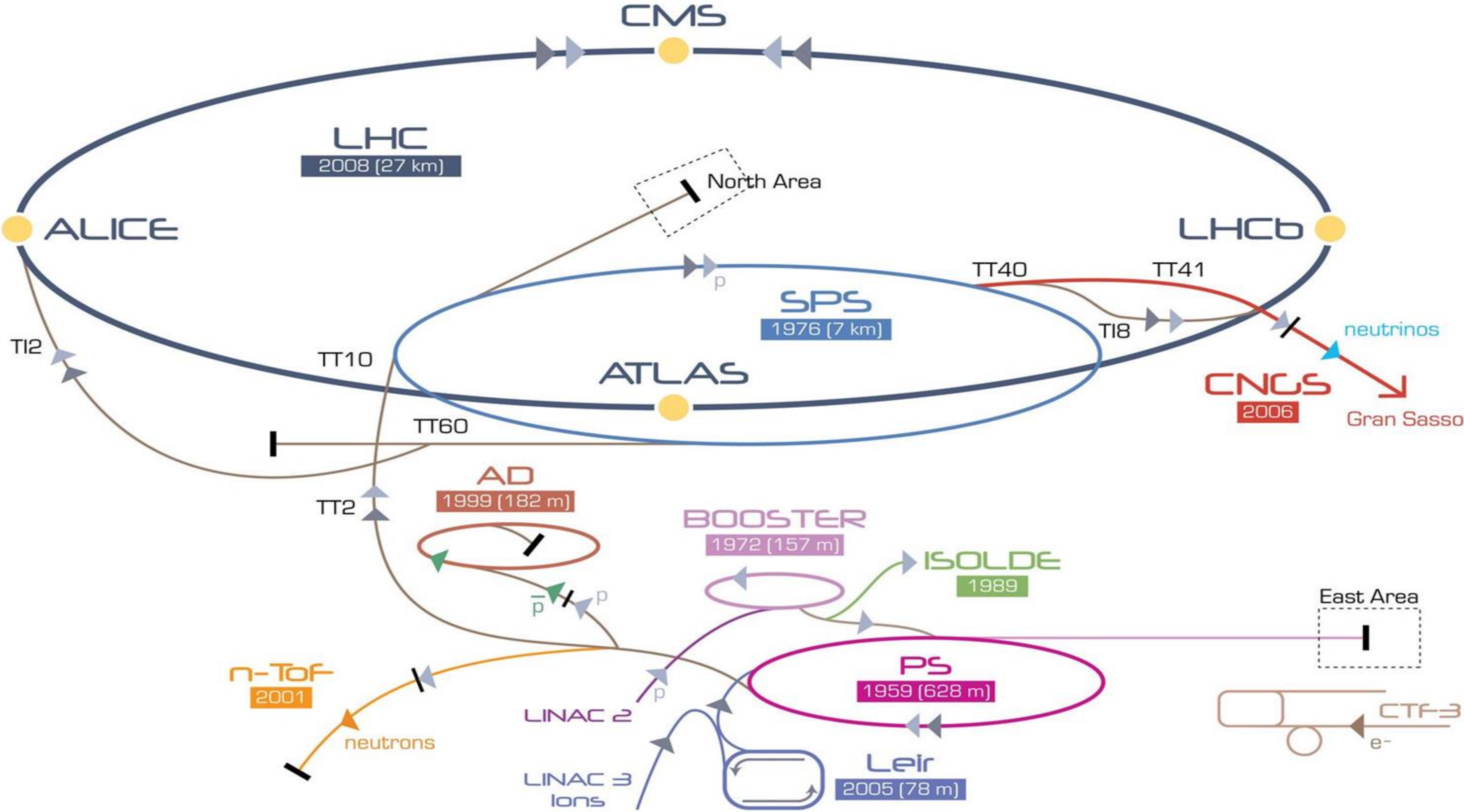
lay the groundwork for for new discoveries

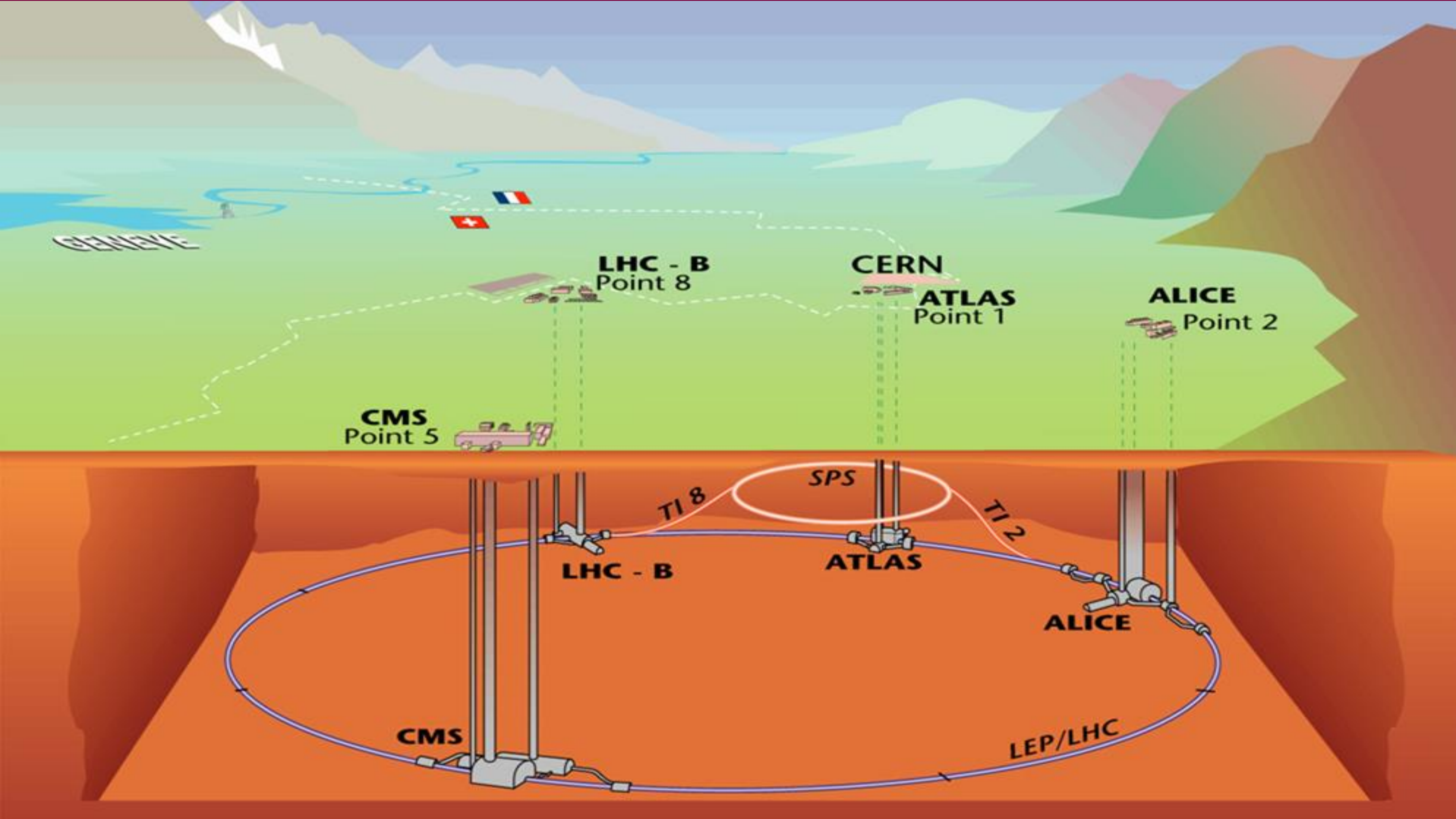


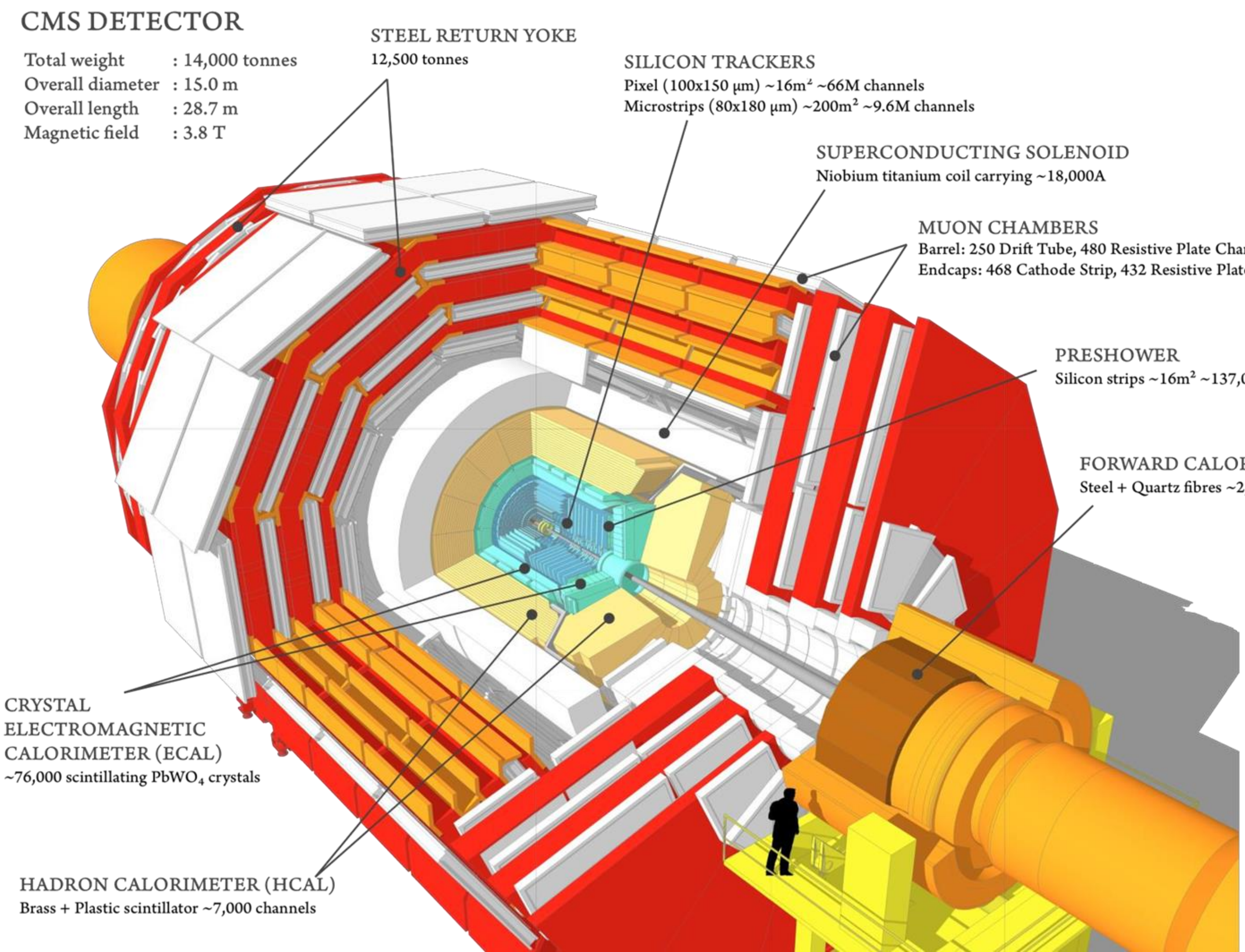
Vilniaus universitetas

- understanding fundamental laws of nature at the smallest scales
- reproduce conditions similar to early after the big bang in the laboratory
- higher energy ⟷ closer in time to big bang


Mt. Blanc
Genève
Lac Léman
LHCb
ATLAS
CERN Meyrin
CERN Prévessin
SPS 7 km
SUISSE
FRANCE
ALICE
CMS
LHC 27 km

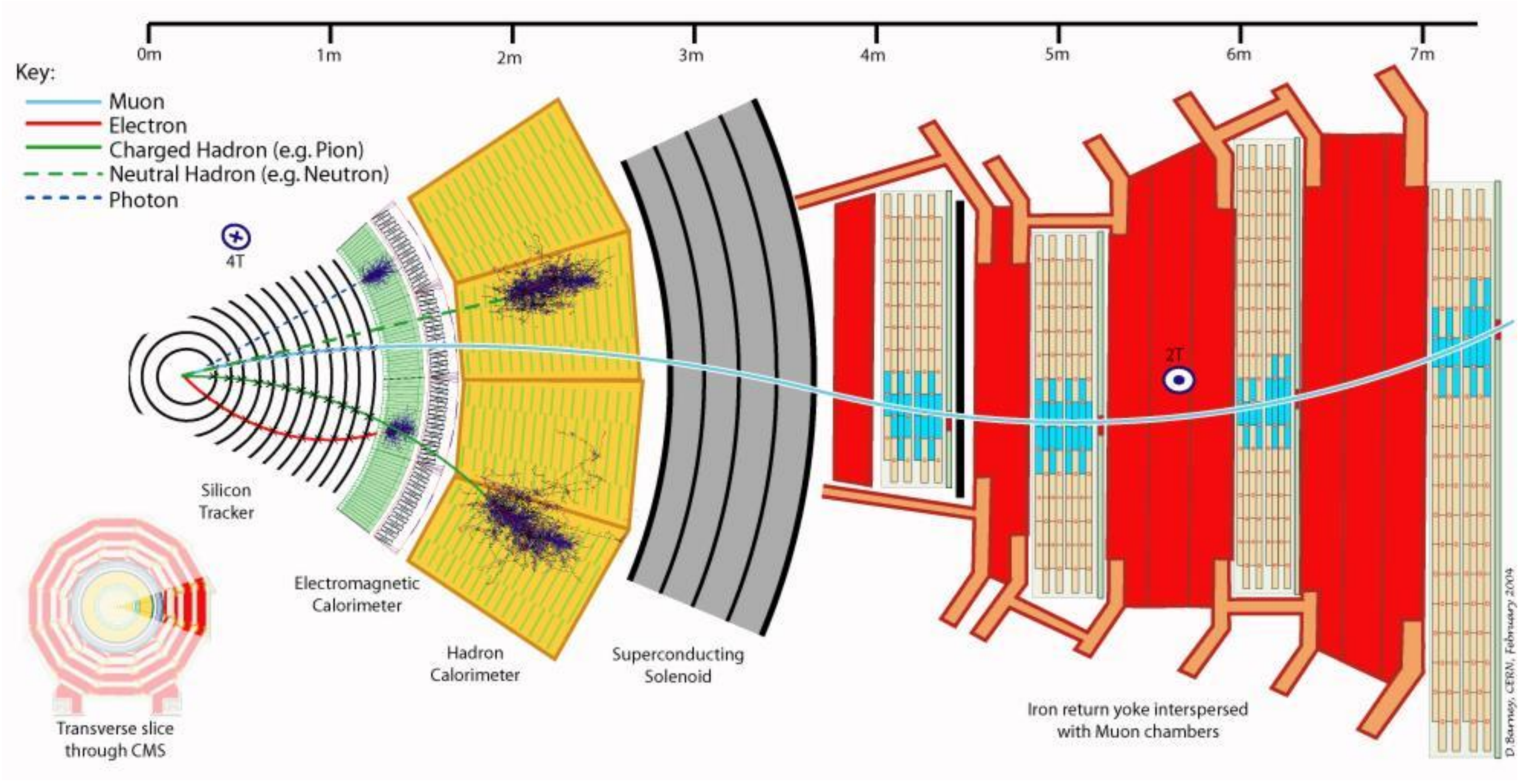

CMS
LHC
2008 (27 km)
North Area
ALICE
LHCb
TT40
TT41
p
SPS
1976 (7 km)
TI8
neutrinos
TI2
TT10
ATLAS
CNGS
2006
Gran Sasso
TT60
AD
1999 (182 m)
TT2
BOOSTER
1972 (157 m)
ISOLDE
1989
p̄
p
East Area
n-ToF
2001
PS
1959 (628 m)
p
LINAC 2
CTF3
e–
neutrons
LINAC 3
Ions
Leir
2005 (78 m)


GENEVE
LHC - B
Point 8
CERN
ATLAS
Point 1
ALICE
Point 2
CMS
Point 5
SPS
TI 8
TI 2
LHC - B
ATLAS
ALICE
CMS
LEP/LHC



Vilniaus universitetas

Vilniaus universitetas



Vilniaus universitetas

https://videos.cern.ch/record/1406325


ME&RE+1/2/16
ME&RE+1/2/24


ME&RE+1/2/16
ME&RE+1/3/18
ME&RE+1/2/18
ME&RE+1/3/19
ME&RE+1/2/19
ME&RE+1/2/20
ME&RE+1/3/20
ME&RE+1/2/21
ME&RE+1/2/22
ME&RE+1/3/21
ME&RE+1/2/24
ME&RE+1/3/22

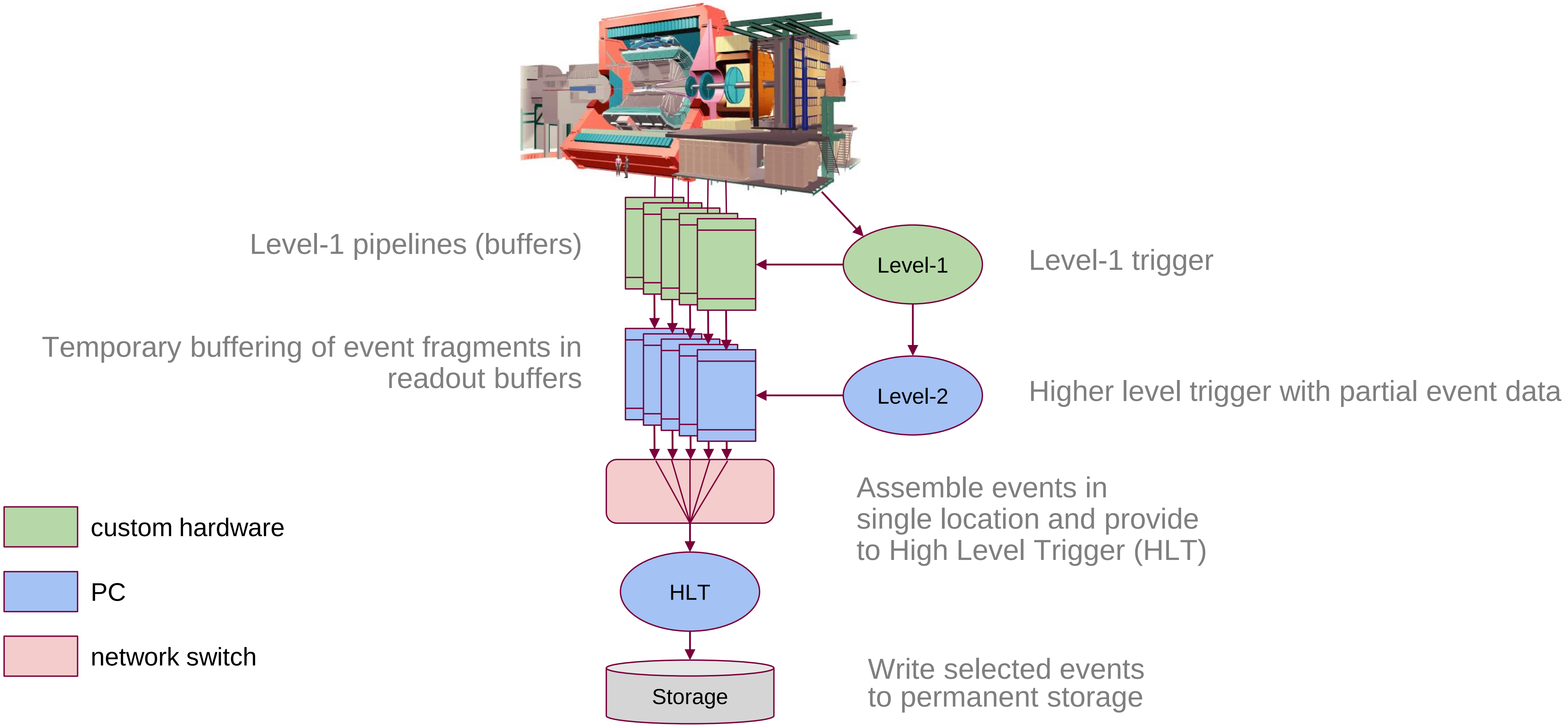
Data Readout System for a Large HEP Experiment

- ~600M collisions / sec, ~1:1M is interesting
- LHC detector scale
  - The number of channels:
    - for LHC experiments $O(10^7)$ channels
    - a (digitized) channel can be between 1 and 14 bits
  - The rate:
    - for LHC experiments everything happens at 40.08 MHz, the LHC bunch crossing frequency
    - This corresponds to 24.9500998 ns or ~25 ns
- HEP experiments consist of many different sub-detectors:
  - tracking, calorimetry, particle-ID, muon-detectors

Data Readout System for a Large HEP Experiment

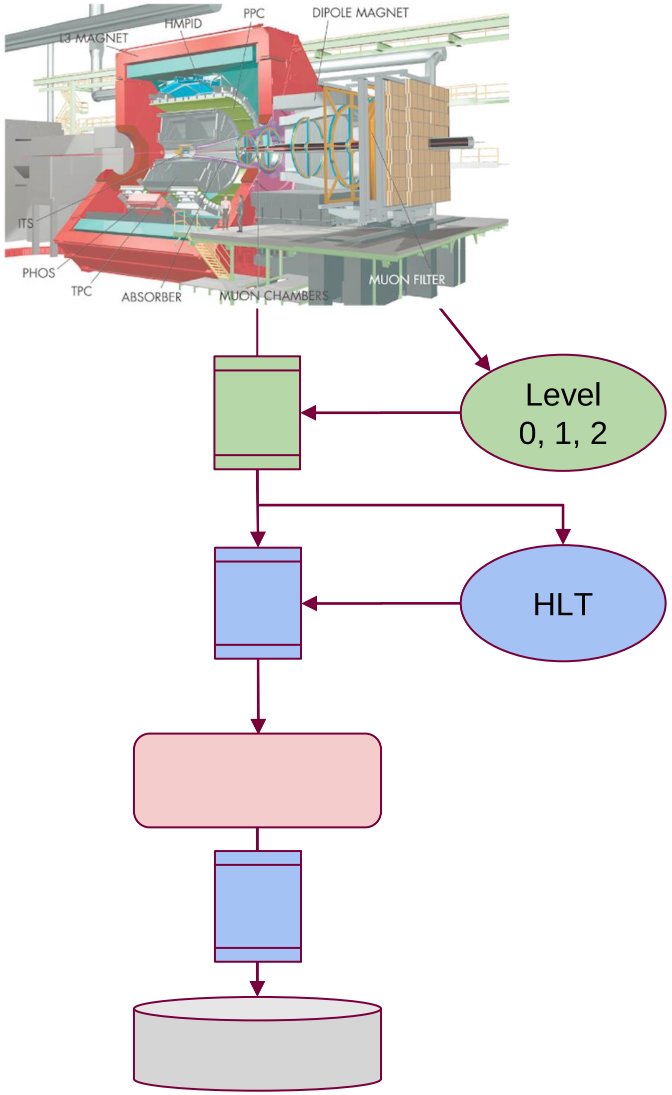
- A selection mechanism (“**trigger**”)
- Electronic readout of the sensors of the detectors (“**front-end electronics**”)
- A system to keep all those things in sync (“**clock**”)
- A system to collect the selected data (“**DAQ**”)
- A Control System to configure, control and monitor the entire DAQ (“**DCS**”)
- A system to monitor data quality and integrity (“**DQM**”)

Data Acquisition “Standard Model”



Vilniaus universitetas

Data Acquisition at ALICE



Front-End pipelines

500 Hz

Higher level trigger

100 Hz

Event Builder

100 Hz

Event buffer

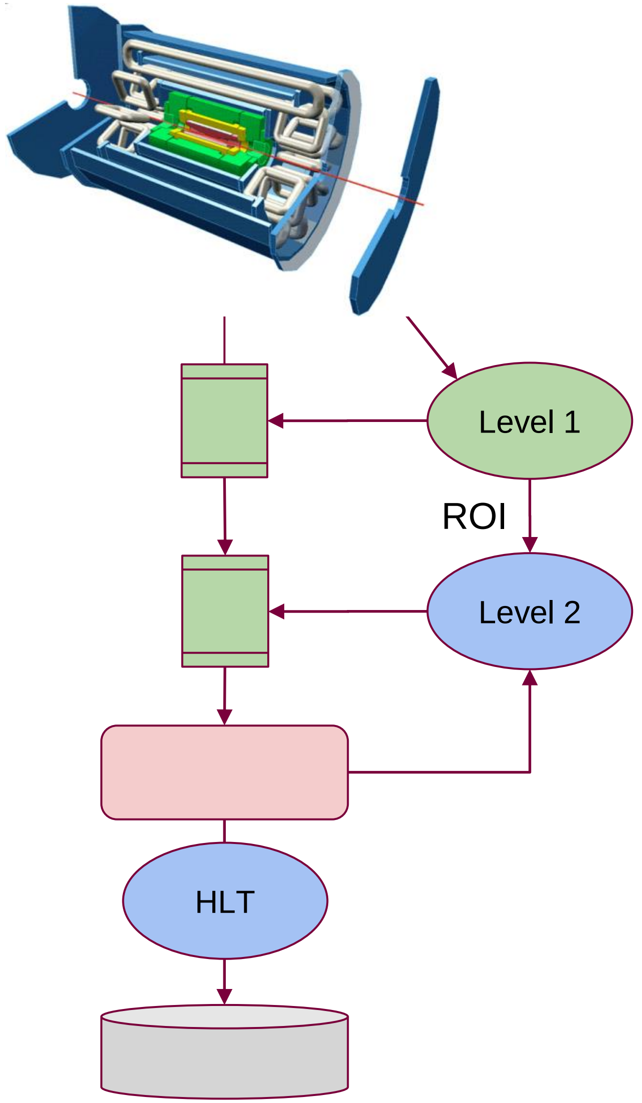
custom hardware

PC

network switch

Vilniaus universitetas

Data Acquisition at ATLAS



40 MHz

Front-End pipelines

100 kHz

Readout buffer

3kHz (L2)

Event Builder

High Level Trigger

200 Hz

custom hardware

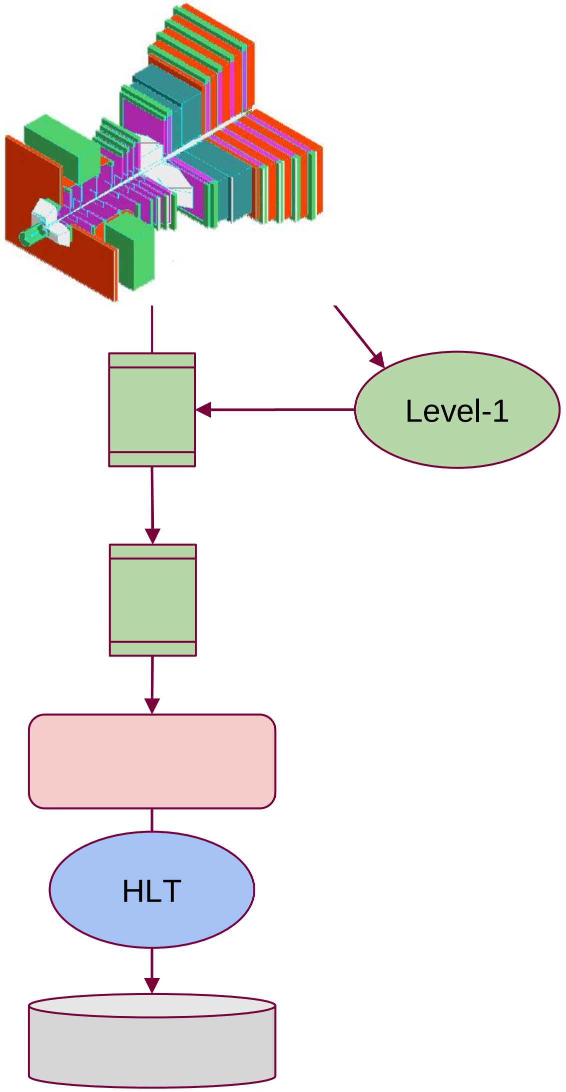
PC

network switch

ROI Regions Of Interest

Vilniaus universitetas

Data Acquisition at LHCb



10 MHz

Front-End pipelines

1 MHz

Readout Buffer

Event Builder

High Level Trigger

2 kHz

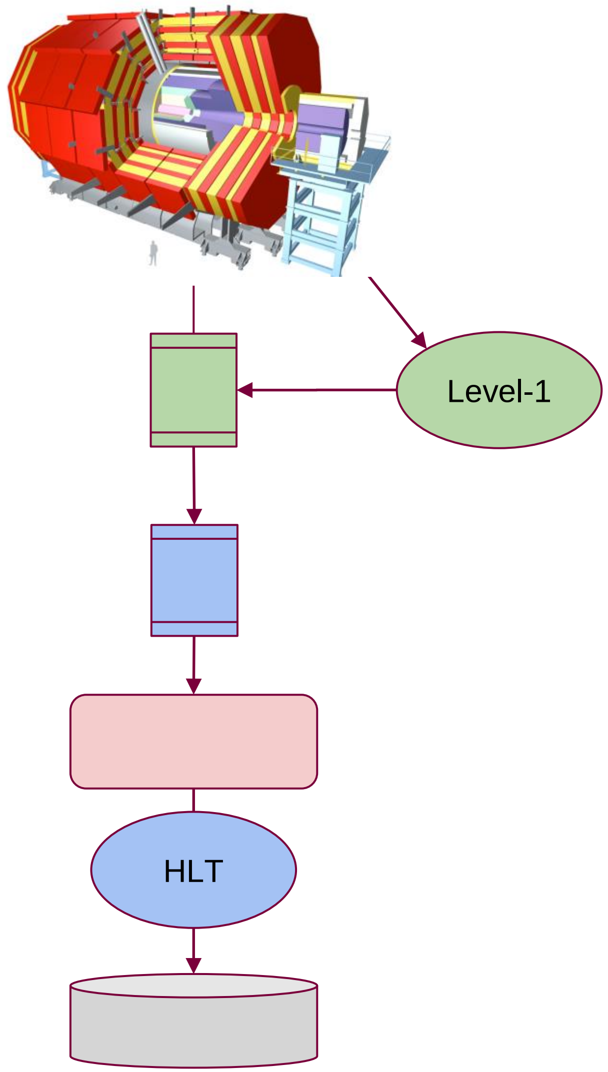
custom hardware

PC

network switch

Vilniaus universitetas

Data Acquisition at CMS



40 MHz

Front-End pipelines

100 kHz

Readout Buffer

Event Builder

High Level Trigger

100 Hz

custom hardware

PC

network switch

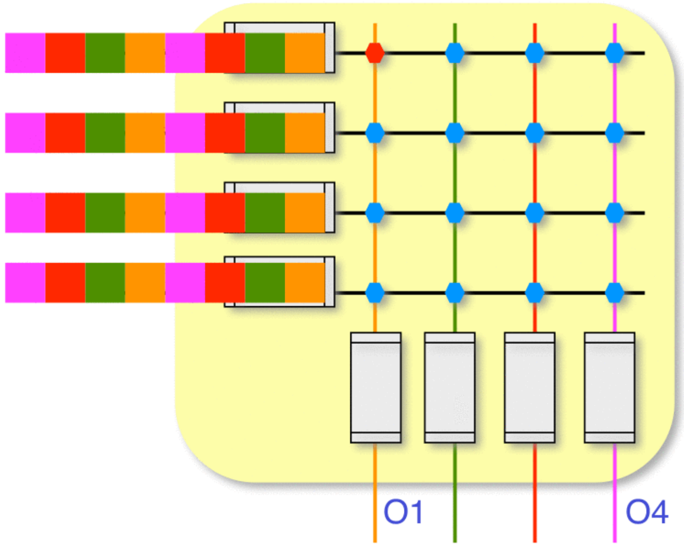
Level-0,1,2,* Trigger

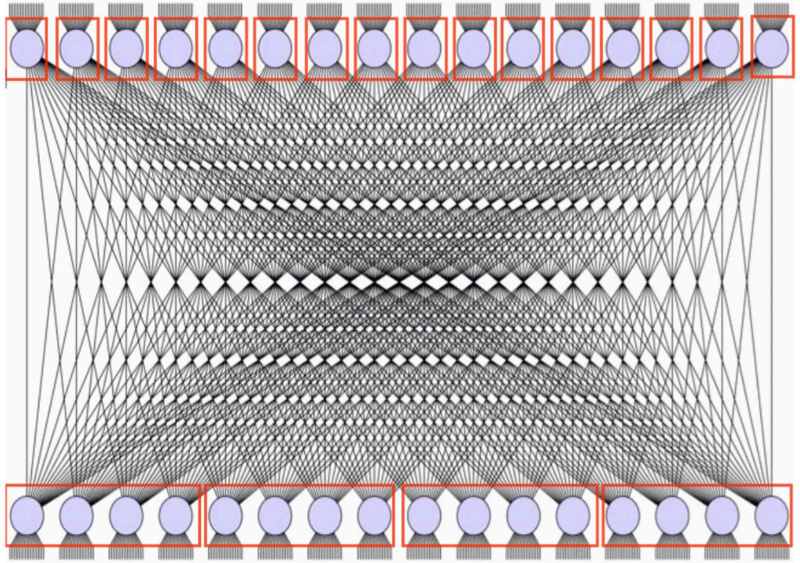
- No (affordable) DAQ system could read out O(10^7) channels at 40 MHz → 400 TBit/s to read out – even assuming binary channels!
- Most of these millions of events per second are totally uninteresting: one Higgs event every 0.02 seconds / Low level triggers must somehow select the more interesting events
- Design principles
  - Millions of channels → try to work as much as possible with “local” information / Keeps number of interconnections low
  - Must be fast: look for “simple” signatures / Keep the good ones, kill the bad ones / Robust, can be implemented in hardware (fast) / to keep buffer sizes under control / every 25 nanoseconds (ns) a new event: have to decide within a few microseconds (μs): trigger-latency

Event Builder

Vilniaus universitetas

- After L1, event data are digitized, pre-processed and tagged with a unique, monotonically increasing number
- The event data are distributed over many read-out boards (“sources”)
- For the next stage of selection, or even simply to write it to tape we have to get the pieces of the event together: **event building**
- L1 rate: 100kHz, event size: 1 Mbyte, No. readout systems: 512





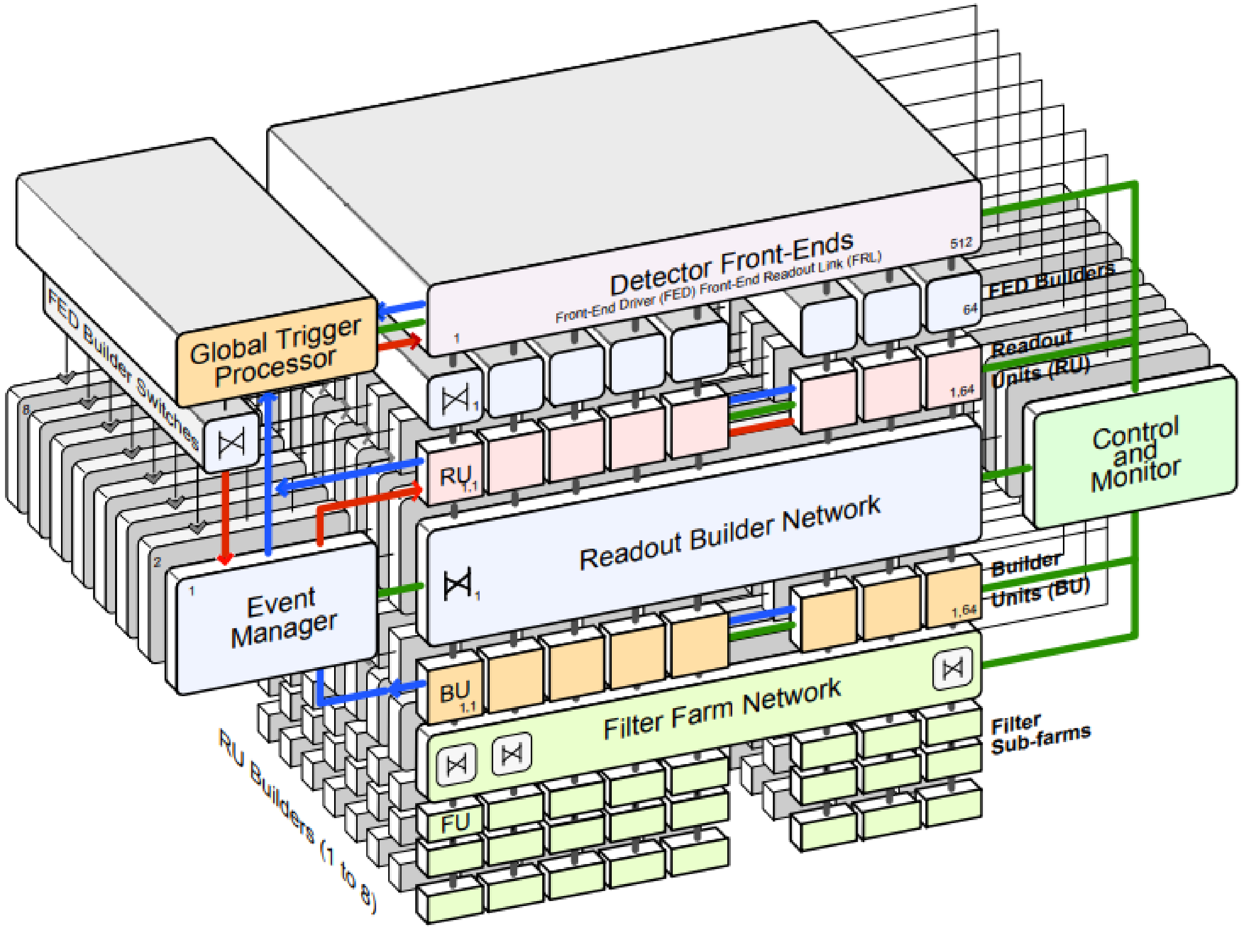
Vilniaus universitetas

High-Level Trigger

- assembly of full collision data on second layer of PCs
- must reduce the rate of selected collisions from 100 kHz to ~ 1 kHz
- ~ 15’000 cores → 150 ms decision time on average
- software of 3.8M C++ and 1.2M python lines of source code
- partial reconstruction of collision data
  - finding clusters of high energy deposit
  - 3D track fitting (Kalman filtering) from 3D and 2D points
  - matching of tracks to clusters

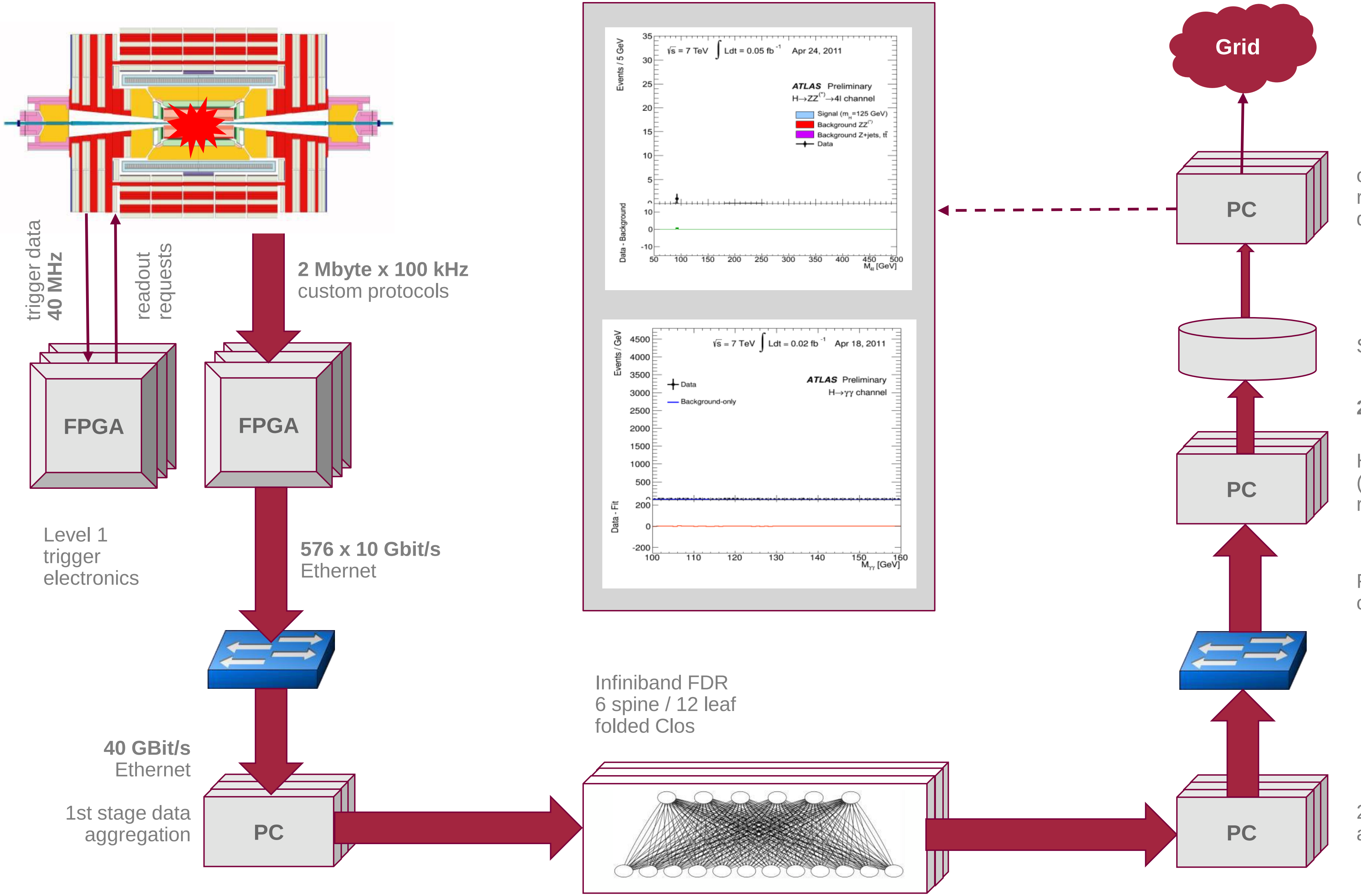
CMS DAQ

Vilniaus universitetas



- Collision rate 40 MHz
- Level-1 Maximum trigger rate 100 kHz
- Average event size ≈ 1 Mbyte
- Event Flow Control ≈ 10^6 Mssg/s
- No. of In-Out units: 512
- Readout bandwidth ≈ 1 Terabit/s
- Event filter computing power ≈ 10^6 SI95
- Data production ≈ Tbyte/day
- No. of PC motherboards ≈ Thousands


BSRT Beam 1
BSRT Beam 2
06:13:43
BEAM SETUP: SQUEEZE
LHC Machine Coordination Planning
Q-FPGA
Q-FPGA



Vilniaus universitetas

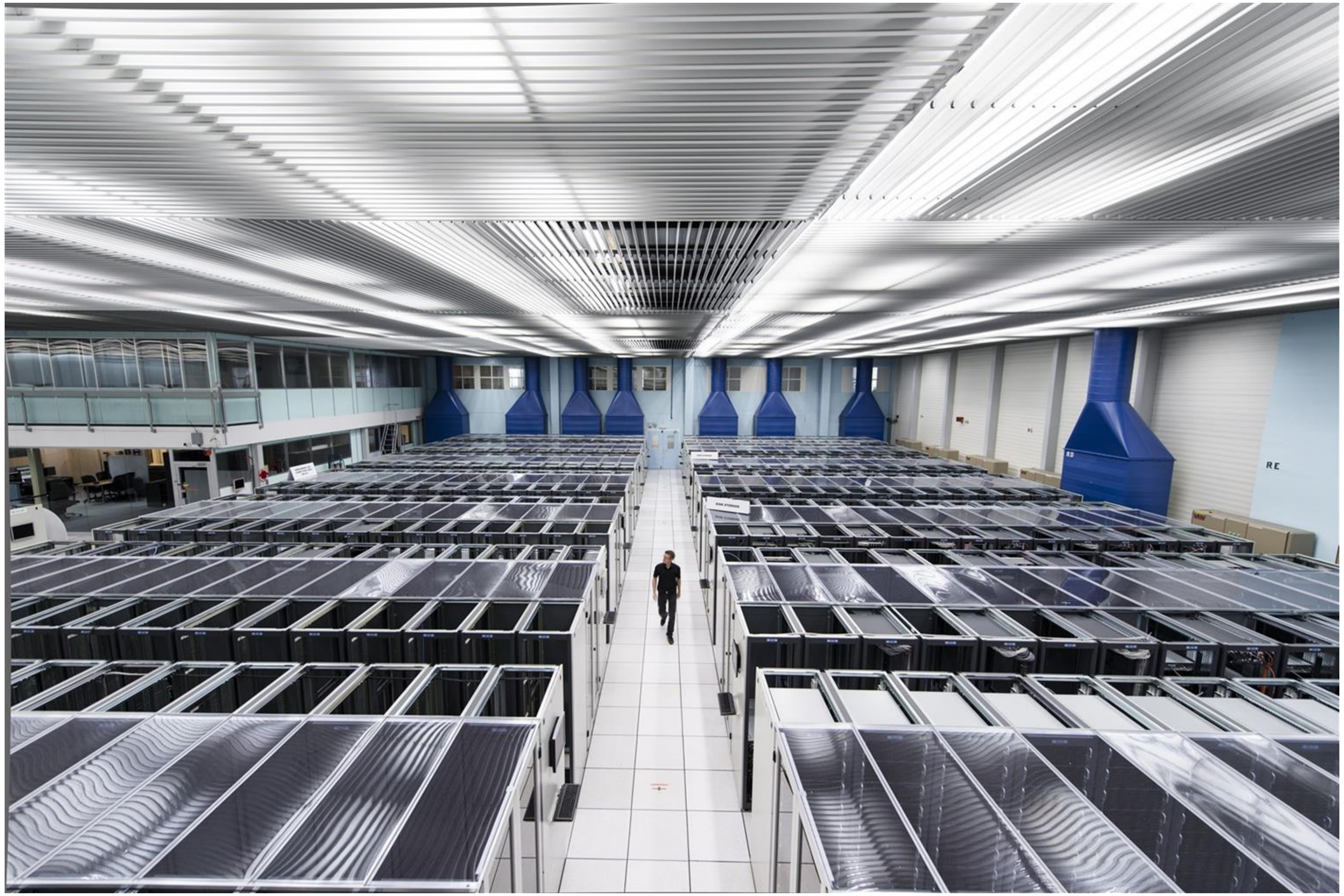
Data Centre

Most of the CERN IT equipment is hosted in the Meyrin Data Centre. However, a second network hub has been inaugurated in 2017 and is located in Prévessin.

About 470 000 processor cores and 11 000 servers run 24/7

90% of resources run a private OpenStack cloud which hosts around 14 000 virtual machines

380 PB of data on tapes, in 2020 has served 2.5 exabyte of physics data, increase of 25 PB/year



Worldwide LHC Computing Grid

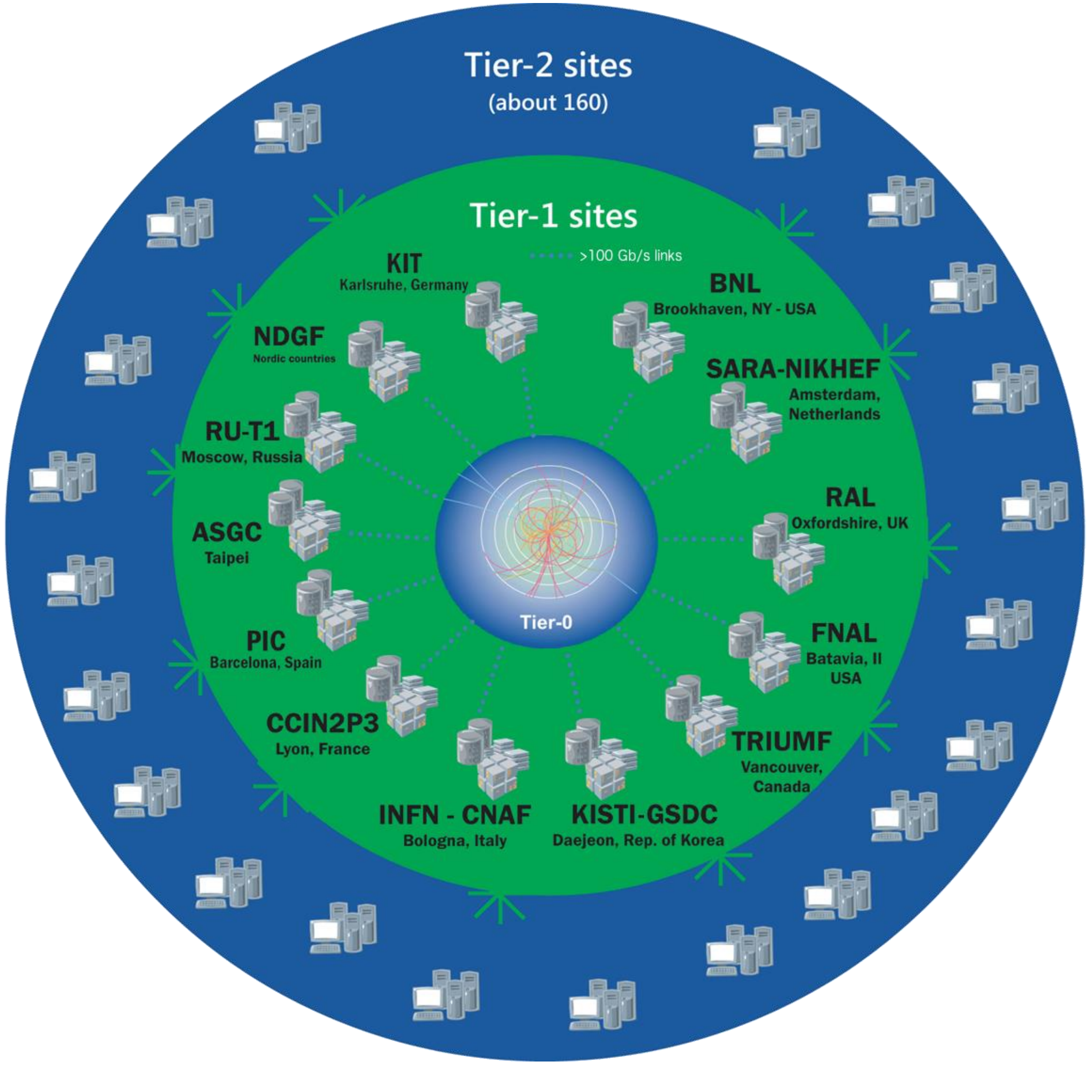
WLCG project is a global collaboration of around 170 computing centres in more than 40 countries

Tier-0: CERN Data Centre, around 20% of the total compute capacity, safe-keeping of the raw data (first copy), first pass reconstruction, distribution of raw data

Tier 1: 13 large computer centres / safe-keeping of a proportional share of raw and reconstructed data, large-scale reprocessing / distribution of data to Tier 2s

Tier 2: around 160 universities and other scientific institutes, handle analysis requirements and share of simulated event production and reconstruction

Tier 3: individual scientists will access these facilities through local computing resources

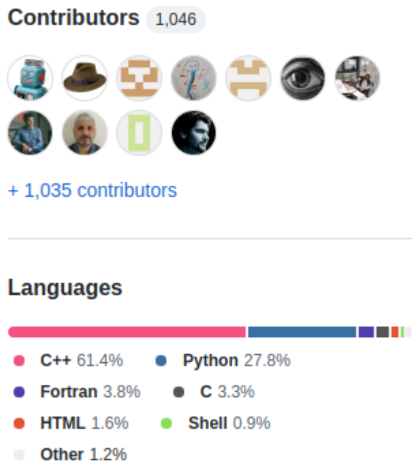


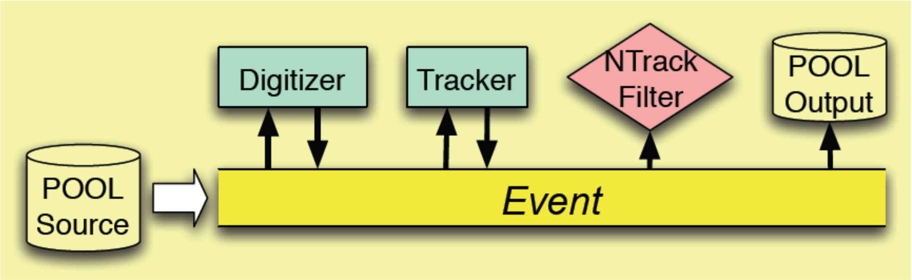
Large Hadron Collider data processing https://www.youtube.com/watch?v=jDC3-QSiLB4

Vilniaus universitetas

CMSSW

- The CMS Software (CMSSW) is a collection of software that the CMS experiment uses in order to acquire, produce, process and even analyze its data
- The program is written in C++ but its configuration is manipulated using the Python language.
- CMSSW is built around a Framework, an Event Data Model (EDM), and Services needed by the simulation, calibration and alignment, and reconstruction modules that process event data so that physicists can perform analysis
- http://cms-sw.github.io/

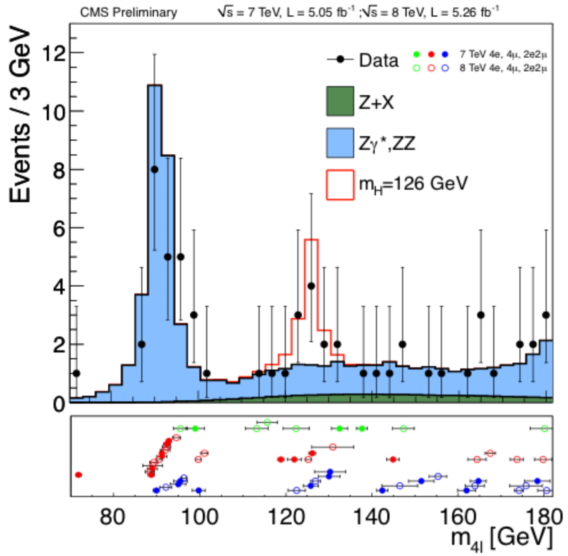




Vilniaus universitetas

ROOT

- An open-source data analysis framework used by high energy physics and others
    - Save and access your data (and any C++ object) in a compressed binary form in a ROOT file
    - Mine data by using powerful mathematical and statistical tools are provided to operate on your data
    - Results can be displayed with histograms, scatter plots, fitting functions and others
    - Use the Cling C++ interpreter for your interactive sessions and to write macros, or you can compile your program to run at full speed
    - ROOT provides a set of bindings in order to seamlessly integrate with existing languages such as Python and R
- https://root.cern/

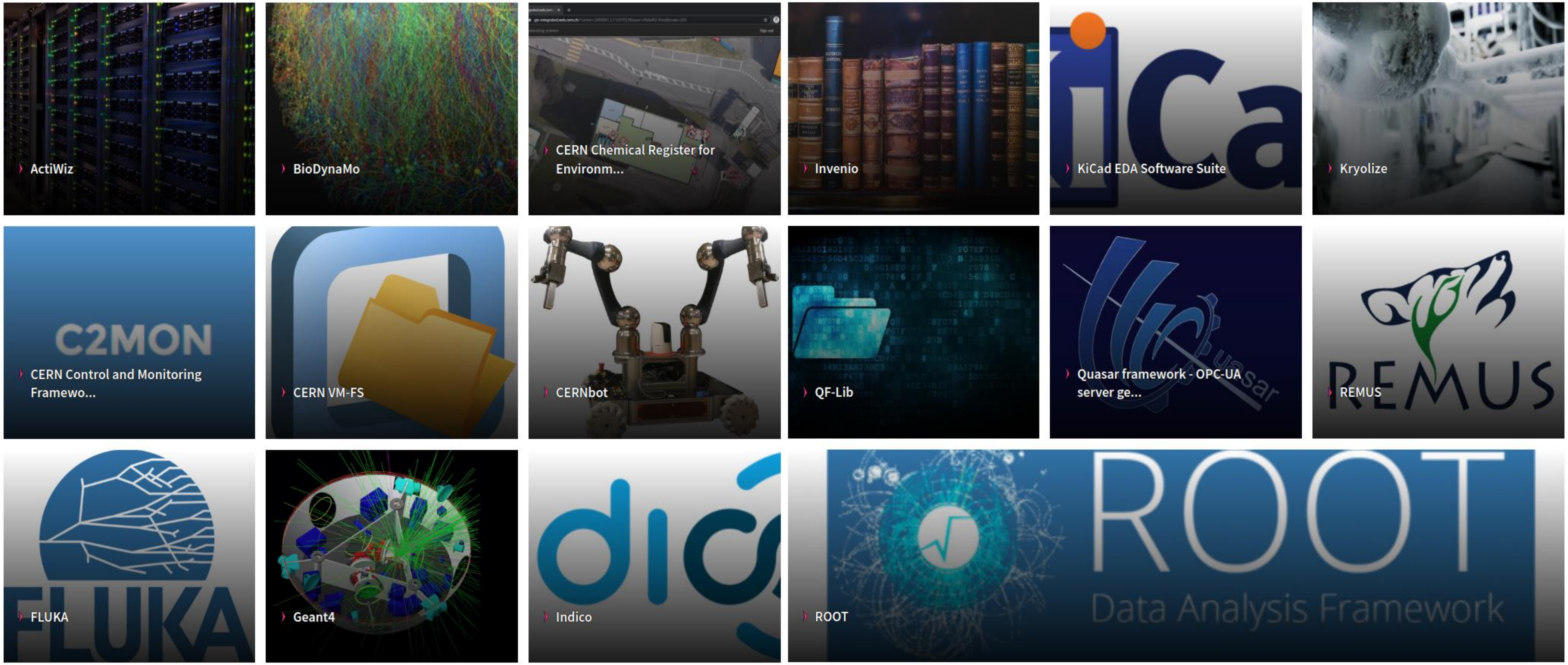


Linux at CERN

- CERN used to be a Red Hat Enterprise Linux customer. But, back in 2004, they worked with Fermilab to build their own Linux distribution called Scientific Linux.
- Eventually they realized that, because they were not modifying the kernel, there was no point in spending time spinning up their own distribution. In 2015, CERN began migrating away from Scientific Linux to CentOS. Scientific Linux is still maintained by a Fermilab, other labs and universities.
- On December 8, 2020, IBM's Red Hat announced the discontinuation of CentOS. CERN turned to alternative - AlmaLinux. Scientific Linux 7, at Fermilab, and CERN CentOS 7, at CERN, will continue to be supported for their remaining life, until June 2024.
- AlmaLinux, a somewhat popular free Linux distribution derived from Red Hat Enterprise Linux (RHEL), received a vote of confidence on Thursday from the European and American science communities.

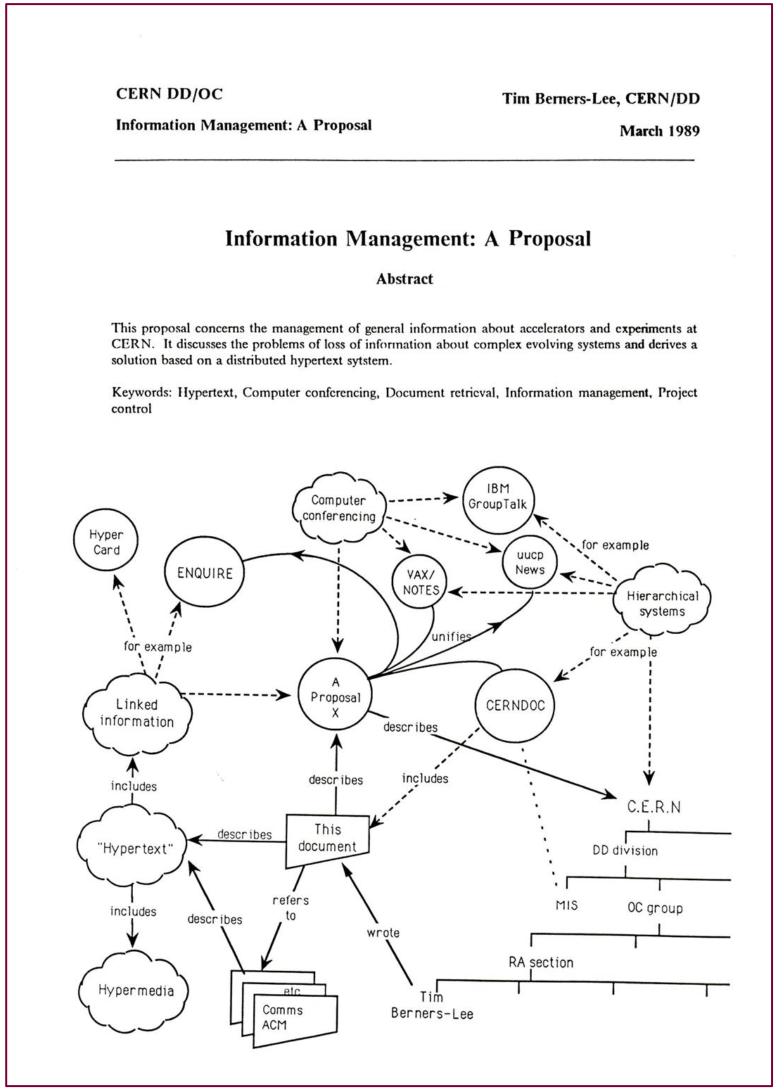
Vilniaus universitetas

CERN ITC Technologies



Vilniaus universitetas

World Wide Web



CERN DD/OC
Information Management: A Proposal

Tim Berners-Lee, CERN/DD
March 1989

Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

VU MIF students at CERN

Vilniaus universitetas

Opportunities for IT/Computing students

- CERN-CMS-Erasmus projects since 2008
- IRIS-HEP projects since 2024

Students contributing to CMS projects in spring/2024

- “CMS e/g Level-1 Trigger: Parallelization and scheduling of a python-based analysis framework using Dask”
- “FPGA NIC for CMS 40 MHz Scouting”
- “Dynamic Web-base Graph Representation of CMSSW modules' dependencies”
- “Fast and efficient data pipelines for training ML models”

Collaboration with University Colleagues including Faculty of Mathematics and Informatics from Vilnius University in Lithuania

This twiki page collects projects in CMS for students coming from Universities wishing to send their students to CERN as a contribution to CMS and as training for the students, for example, Vilnius University in Lithuania:

In 2008 the Faculty of Mathematics and Informatics from Vilnius University in Lithuania established a collaborating with the CMS experiment in giving 3-5 students in spring and summer every year the occasion to work on a specific project in the experiment. The students work on a specific project in the various teams of CMS, typically developing new software which is then used in the collaboration. This way the students get immediate feedback on their work and face the challenges of "real life" feedback in their projects.

IRIS/HEP Fellows Program

The Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP) has been established to meet the software and computing challenges of the planned upgrades to the Large Hadron Collider, the world's most powerful particle accelerator. IRIS-HEP pursues R&D for the software needed to acquire, manage, process and analyze the torrent of data that will be produced by the upgrade accelerator and its detectors as part of the search for discoveries beyond the Standard Model of Particle Physics.

iris-hep.org

Vilniaus universitetas

Thank you!

Valdas Rapševičius
valdas.rapsevicius@mif.vu.lt