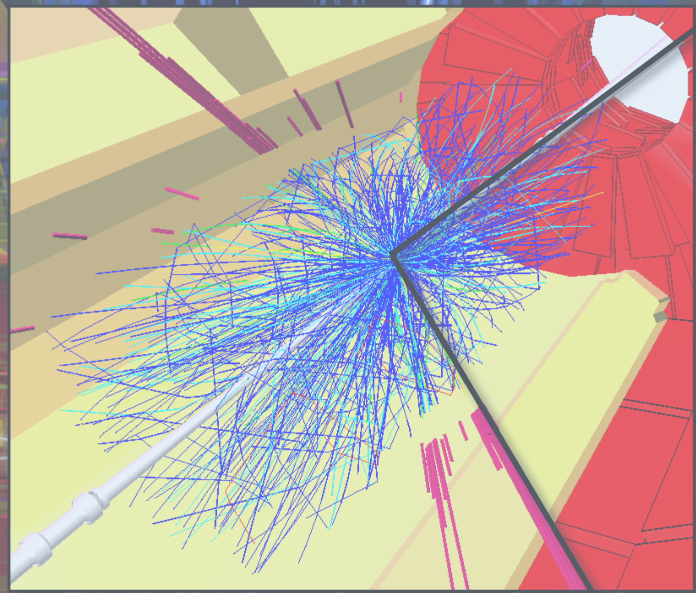
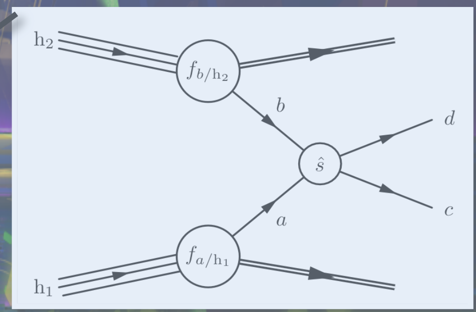
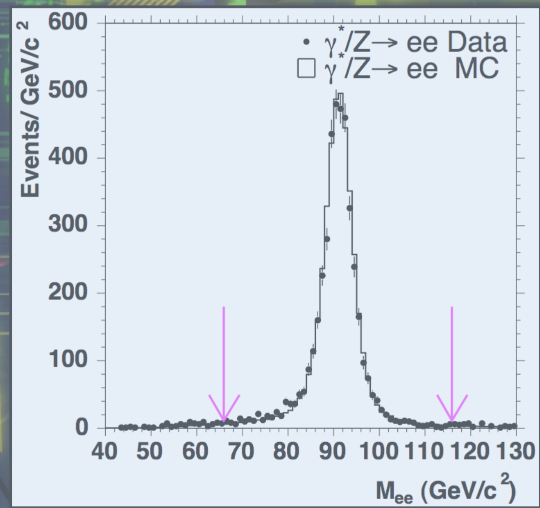
Информационные технологии в физике элементарных частиц







Национальный Исследовательский Ядерный Университет «МИФИ»

март, 2020

Алексей Климентов

Alexei Klimentov

- **Expertise and interests:** High-performance and high-throughput computing, Workload and Data Management Systems, HEP application software, LCF usage for compute-intensive applications

Brief bio: Graduated from MEPhI, PhD in particle physics, Doctor of Science in IT, L3 experiment on Large Electron-Positron Collider, online computing coordinator, AMS experiment (shuttle flight STS91 and International Space Station) SW&Computing Coordinator, Physics Applications Group Head at BNL Physics Department, Leading PI US DOE ASCR and HEP funded project "Next Generation Workload Management System for BigData" (aka BigPanDA, 2012-2019). Over six hundred publications and extensive service on national and international advisory and conference program committees. Founder of ATLAS Distributed Computing and its leader in 2009-2013 (during Large Hadron Collider preparation, commissioning and Run 1). Led several R&Ds in cloud and supercomputing for HEP. Russian Ministry of Education and Science mega-grant award for Leading Foreign Scientists (2014-2018).

Thanks for slides / materials :

- F.Barreiro (UTA)
- K.Bhatia (Google)
- I.Bird (CERN)
- D.Britton (GridPP)
- R.Brun (CERN)
- P.Calafiura (LBNL)
- S.Campana (CERN)
- J.Catmore (U Oslo)
- T.Childers (ANL)
- D.Costanzo (U Sheffield)
- K.De (UTA)
- S.Jha (Rutgers/BNL)
- M.Grigorieva (MSU)
- F.Hemmer (CERN)
- A.Kirianov (NRC KI PNPI)
- K.Kissel (Google)
- T.Maeno (BNL)
- R.Mashinistov (UTA)
- S.McKee (U Michigan)
- D.Oleynik (UTA)
- S.Panitkin (BNL)
- J.Shiers (CERN)
- T.Schulthess (ETH)
- M.Schulz (CERN)
- F.Schurmann (EPFL)
- V.Tsulaia (LBNL)
- A.Undrus (BNL)
- V.Voevodin (MSU)
- J.Wells (ORNL)
- T.Wenaus (BNL)
- A.Zarochentsev (SPbSU)

Lectures drew on presentations, discussions, comments, input from many. Thanks to all, including those I've missed

Disclaimer : This lectures will have a "slight" bias towards ATLAS experiment @ LHC and CERN computing

«мой» неполный список аббревиатур (aka TLA)

AMI	ATLAS Metada Interface	**AMGA**	meta-data and GANGA system
AOD	Analysis Object Data	**API**	Application Program Interface
ARDA	Architectural Roadmap towards Distributed Analysis	**ATCOM**	ATLAS task request package
ATHENA	ATLAS interactive analysis framework	**ATLAS**	A Toroidal LHC ApparatuS
CA	certificate authority	**CASTOR**	CERN "home-made" storage system
CAT	Control a Task (AMI package)	**CDR**	Central Data Recording
COC	Computing Operations and Coordination	**COOL**	Conditions DB project
CONDOR	batch system	**CORAL**	LCG project
DB	Data Base	**DAQ**	Data Acquisition system
DC	Data Challenge (can be 1 to 4)	**DDM**	Distributed Data Management
DIAL	Distributed Interactive Analysis	**DN**	Distinguished Name. (grid-proxy-info command)
DQ	Don Quijote, roman by Miguel cervantes de Saavedra	**DQ2**	ATLAS DDM system
DVS	Detector Verification System	**EOWYN**	ATLAS jobs executor (one of many)
ESD	Event Summary Data	**FTS**	File Transfer Services
GRID	just grid, etc	**GUI**	Graphic User Interface
GUID	Grid Universal ID	**HPSS**	Hierarchical Storage System
IS	Information System	**LEXOR**	LCG Executor
LFC	Local File Catalog		
LFN	Logical File Name	**LHCG**	LHC Computing GRID
LSF	batch system	**MonAlisa**	GRID monitoring package
mySQL	RDBMS	**NFS**	Network File System
NG	Nordic Grid	**OKS**	Obect-oriented Kernel System
ORACLE	RDBMS	**OSG**	Open Science Grid, also GRID flavour
PANDA	bear, also ATLAS Production system	**PFN**	Physical File Name
POOL	LCG prject	**RAT**	Request a Task (Task Request Inerface)
RDBMS	Relational DB managementsystem	**RLS**	Replica Location Service
SC	Service Challenge	**3D**	Distributed Database Deployomen
SQL	Simple Query Language	**SE**	Storage Element
SEAL	Shared Environment for Applications	**SFO**	SubFarm Output
SRM	Site Resource Manager	**SURL**	Site URL
SWING	ATLAS SW Integration Working Group	**TDAQ**	Trigger and DAQ system
TRF	Transformation	**Twiki**	Web pages editor
URL	Universal Resource Locator	**VO**	Virtual Organization
VOMS	Virtual Organization Memebership Service	**VUID**	Virtual Universal ID
WMS	Workload Management System		

Outline 1/2

- Computing in High Energy and Nuclear Physics (HENP)
  - ✓ data acquisition, analysis and processing from Hollerith machine to LHC Computing
  - ✓ WWW, mainframes, workstations, personal computers
  - ✓ Programming languages, libraries, SW tools
  - ✓ LHC Computing, looking at the (Big)Data
  - ✓ Physics data flow
- From RAW data to Physics Results
  - ✓ what is data?
  - ✓ data processing and analysis chain
  - ✓ DAQ chain
  - ✓ data simulation and reconstruction goals
  - ✓ data quality
- LHC Computing Model
  - Distributed computing and Grid paradigm
    - Grid key components
    - MONARC model
    - “mesh” computing model
  - Wide Area Network (WAN)

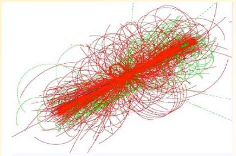
Outline 2/2

- ✓ SW&Computing challenges in High-Luminosity LHC era. Big data at LHC
  - Роль суперкомпьютеров для научной программы физики элементарных частиц
  - Интеграция грид, суперкомпьтеров и ресурсов «облачных вычислений»
  - Intel x86 и новые вычислительные архитектуры
  - Pioneering Research & Development projects in HENP Computing
  - from HENP applications to bioinformatics and computational science
- ✓ Системы управления данными физического эксперимента
  - Системы управления загрузкой
  - Системы для обработки и анализа данных.
  - Системы управления данными физического эксперимента
    - передача данных
    - «популярность данных»
  - Системы мониторинга
  - Эволюция системы управления потоками научных заданий. Новые виды потоков заданий в эпоху суперБак
- ✓ Operations Intelligence и методы визуализации научных данных М.А.Григорьева

Экзамен апрель 2020

Our Task

Reality



We use experiments to inquire about what "reality" (nature) does

We intend to fill this gap

Theory

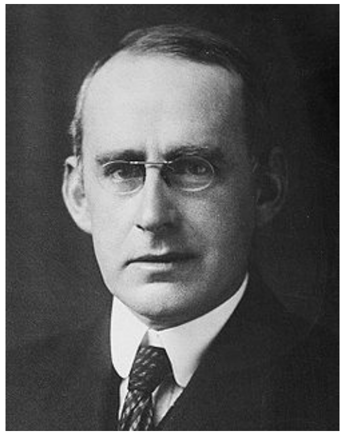
$$S = \mathrm{i} \int \mathrm{d}^4 x \, \mathcal{L}(x)$$

ATLAS/CMS Physics Goals

- Explore high energy frontier of particle physics
- Search for new physics
  - Higgs boson and its properties
  - Physics beyond Standard Model – SUSY, Dark Matter, extra dimensions, Dark Energy, etc
- Precision measurements of Standard Model parameters

The goal is to understand in the most general; that's usually also the simplest.

- A. Eddington



август 2003. Тибет, монастырь Ганден

Data Acquisition

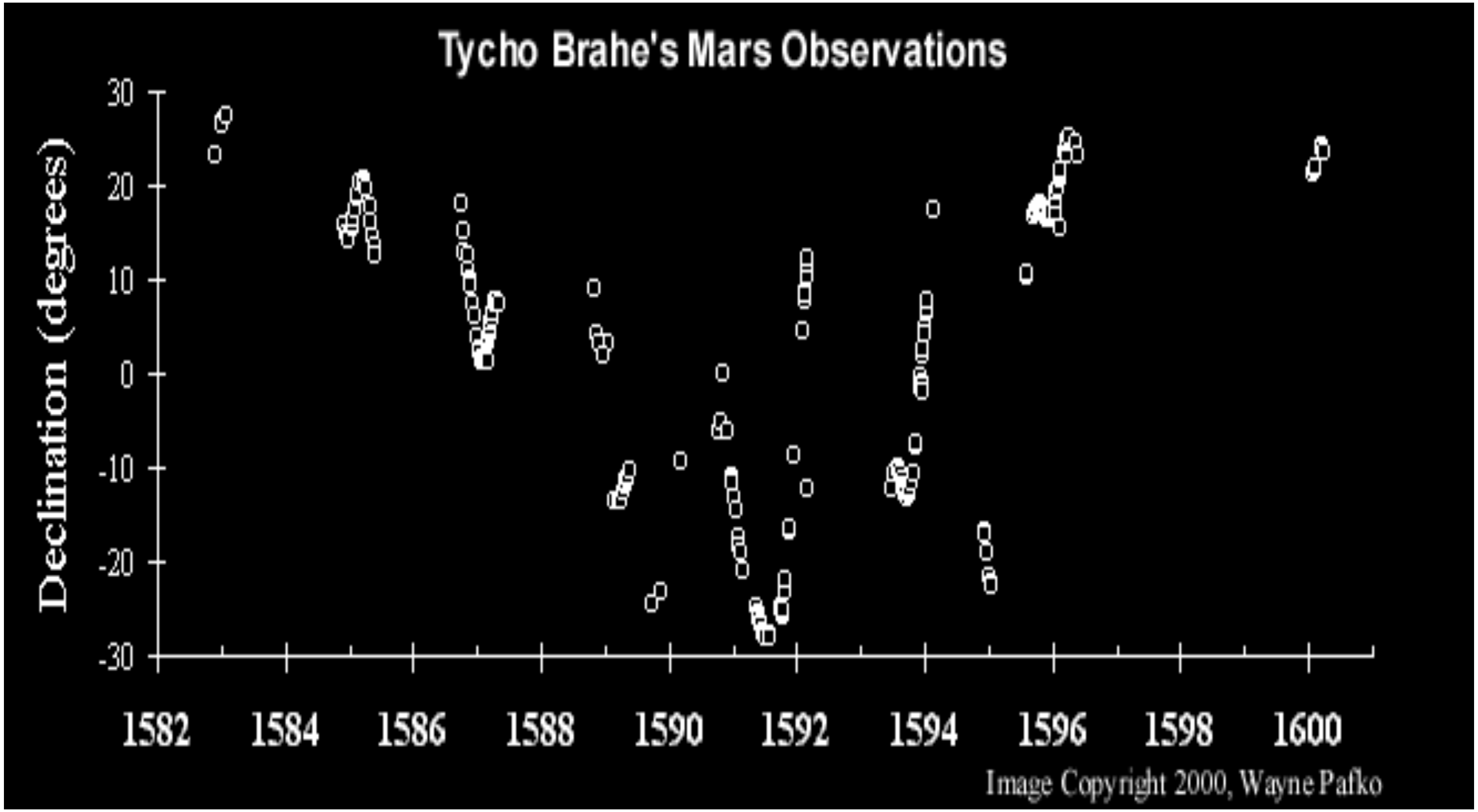
Tycho Brahe and the Orbit of Mars

I've studied all available charts of the planets and stars and none of them match the others. There are just as many measurements and methods as there are astronomers and all of them disagree. What's needed is a long term project with the aim of mapping the heavens conducted from a single location over a period of several years.

Tycho Brahe, 1563 (age 17).

- First measurement campaign
- Systematic data acquisition
  - Controlled conditions (same time of the day and month)
  - Careful observation of boundary conditions (weather, light conditions etc…) - important for data quality / systematic uncertainties

The First Systematic Data Acquisition



- Data acquired over 18 years, normally every month
- Each measurement lasted at least 1 hr with the naked eye
- Red line (only in the animated version) shows comparison with modern theory

Data Analysis

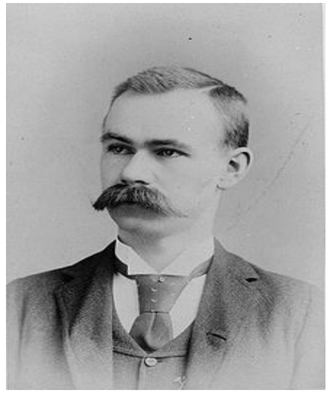
- Tycho did not do the correct analysis of the Mars data, this was done by Johannes Kepler (1571-1630), eventually paving the way for Isaac Newton theory of universal gravitation

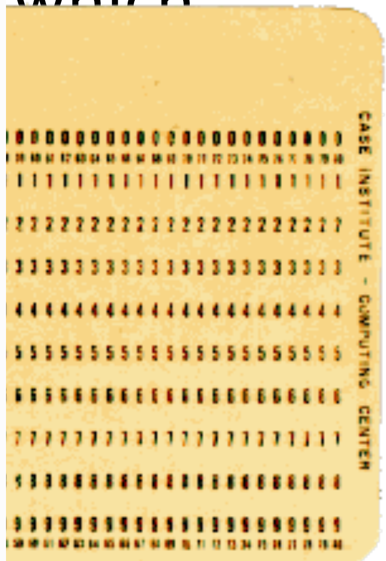
Data Processing

A… .

- Accordin… d in late
  1800s wi… a
  … to
  …ces of
- … which
-

[Jacquard-card Making.]





Hollerith’s Successes

- In 1890 Hollerith founded a company called the Tabulating Machine Company.
- In 1911, his company merged with two other companies to create the Computing-Tabulating-Recording Company.
- Under the direction of Thomas Watson Sr, CTR would change its name in 1924 to International Business Machines. Hollerith's machine would provide the basis for IBM's success and make him the father of information processing.

Thomas Watson Jr

I think there is a world market for maybe five computers

- 1943

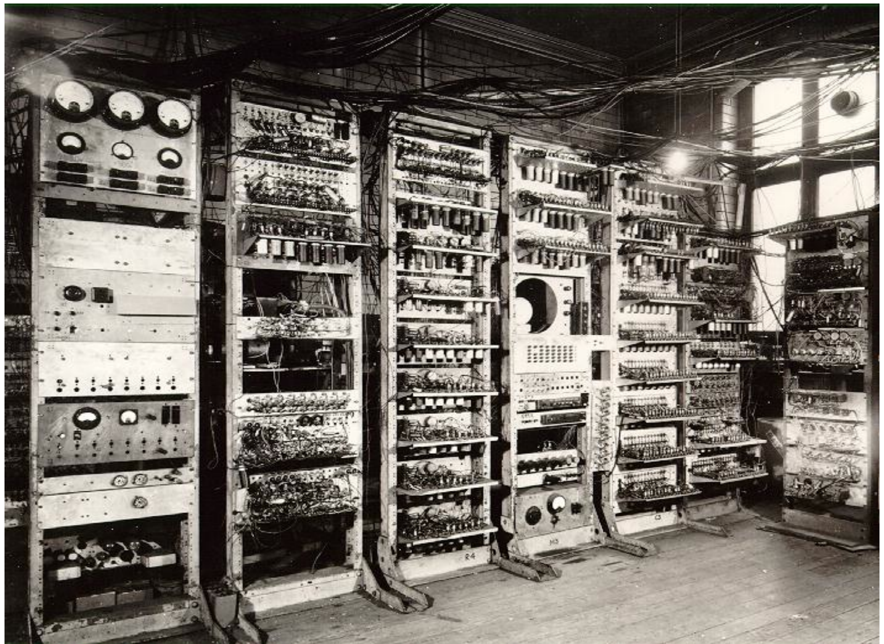
Name	First operational	Numeral system	Computing mechanism	Programming	Turing complete
Zuse Z3 (Germany)	May 1941	Binary floating point	Electro-mechanical	Program-controlled by punched 35 mm film stock (but no conditional branch)	Yes (1998)
Atanasoff–Berry Computer (US)	1942	Binary	Electronic	Not programmable—single purpose	No
Colossus Mark 1 (UK)	February 1944	Binary	Electronic	Program-controlled by patch cables and switches	No
Harvard Mark I – IBM ASCC (US)	May 1944	Decimal	Electro-mechanical	Program-controlled by 24-channel punched paper tape (but no conditional branch)	No
Colossus Mark 2 (UK)	June 1944	Binary	Electronic	Program-controlled by patch cables and switches	No
Zuse Z4 (Germany)	March 1945	Binary floating point	Electro-mechanical	Program-controlled by punched 35 mm film stock	Yes

Computers: Predictions

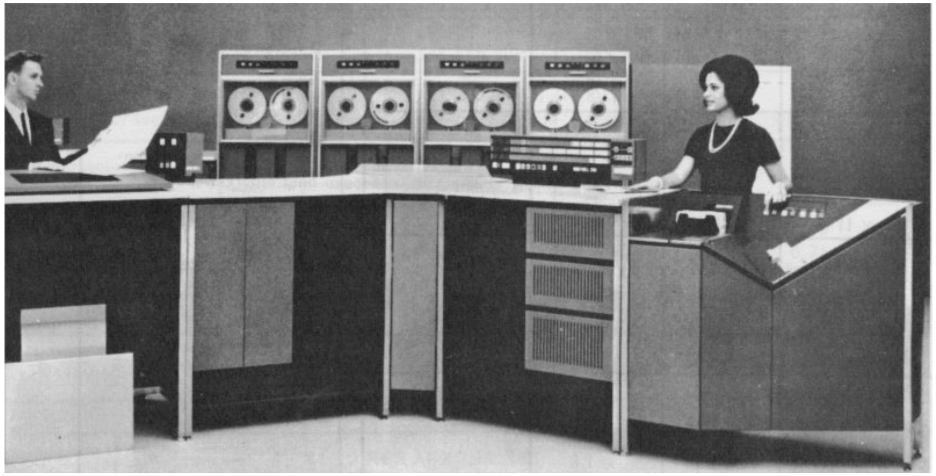
“Computers in the future may weight no more than 1.5 tons”

– Popular mechanics, 1949

The first electronic computer was named Colossus (~1944) and weighed approximately one ton



Hollerith's Successes. Cont'd



- Early 60. Computer services company in industrial Yorkshire, UK.
- Computer: Honeywell H200

Input device

Network device

БЭСМ-6. 1 MIPS. 1966

1960

- At the beginning of the computer era, the Soviet Union competed on an equal footing with the world's leading computer powers
- A masterpiece of computer engineering, in which many revolutionary solutions were realized. The machine was produced for 17 years 1966-1983 (about 450 machines were manufactured) and survived three generations of computer technology. The last copy of the legend worked in the Training Center of the Navy near St. Petersburg.

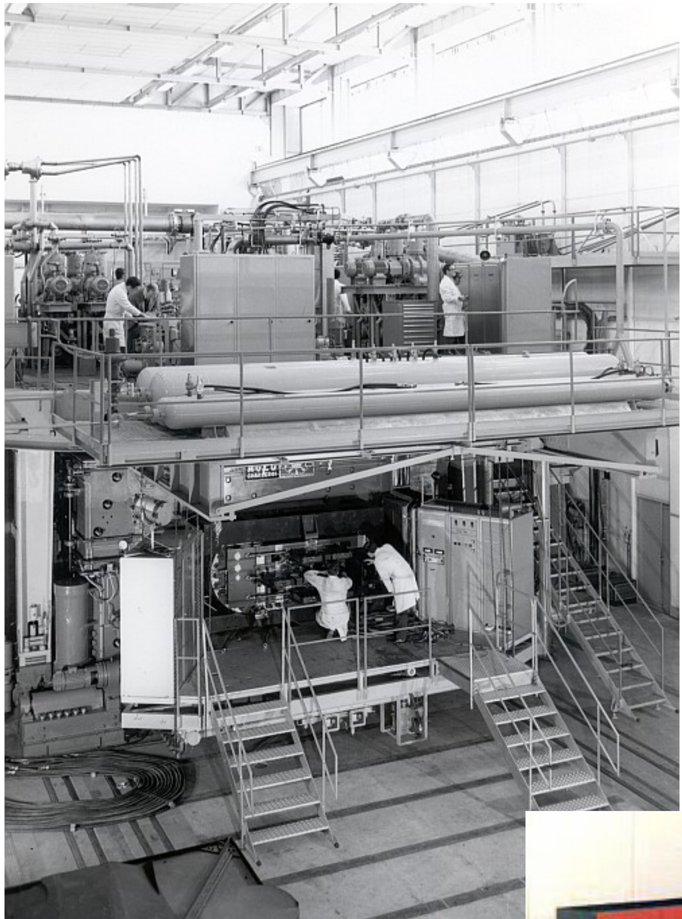
Ken Olson on the PC

- "There is no reason for any individual to have a computer in his home."
  - Ken Olson, president, Digital Equipment Corporation, 1977.
- Ironic that DEC was subsequently taken over by COMPAQ…
- … and COMPAQ was taken over by HP in Jan 2002.

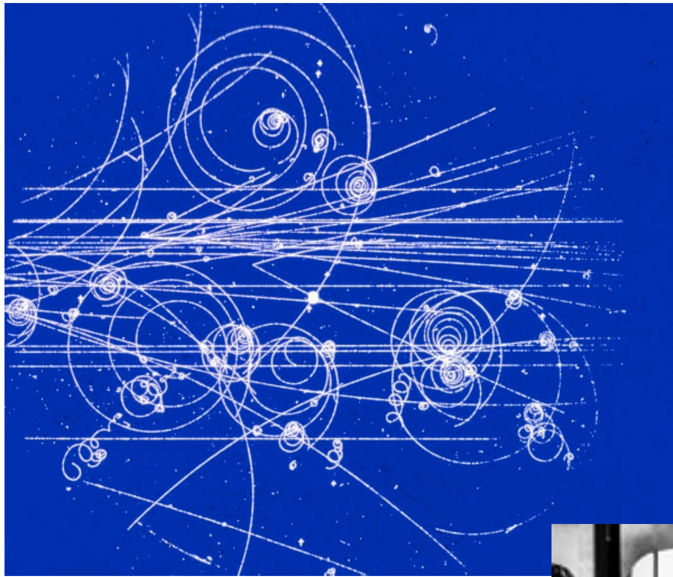
DEC PDP-8/S : ~0.03 MIPS
One of the main processing computers for bubble chamber experiments

DEC – Digital Equipment Corporation
Major models : LSI, PDP 11, VAX

CERN 2M Hydrogen Bubble Chamber



100,000+ Pictures



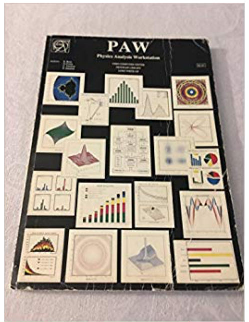
The "trigger" : data filtering

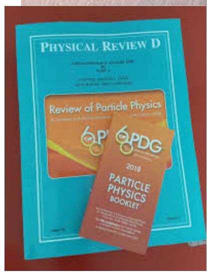
IBM 360/44
Offline Analysis
~0.09 MIPS

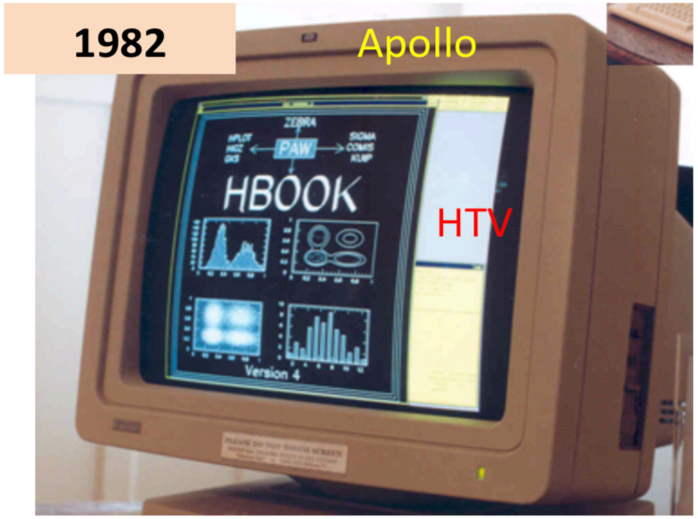
MIPS – Million Instructions per second

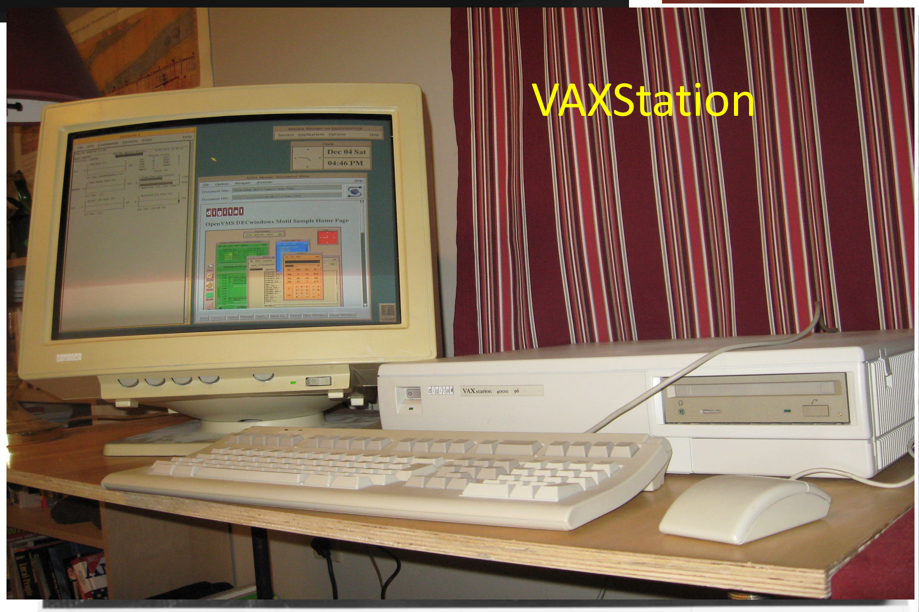
Mainframes & Workstations

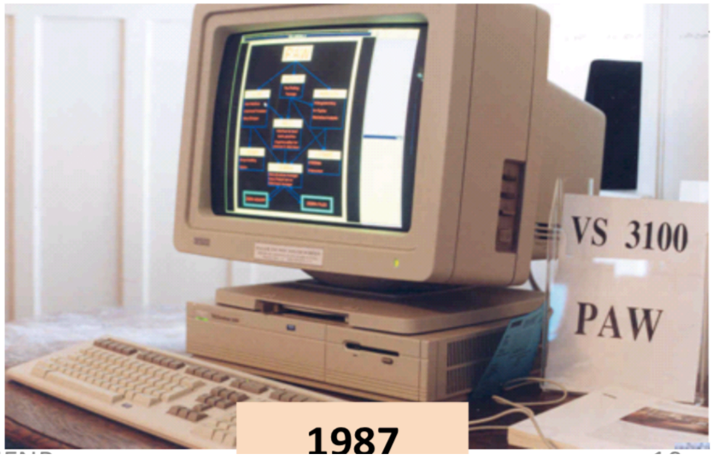
1987 : PAW – Physics Analysis Workstation





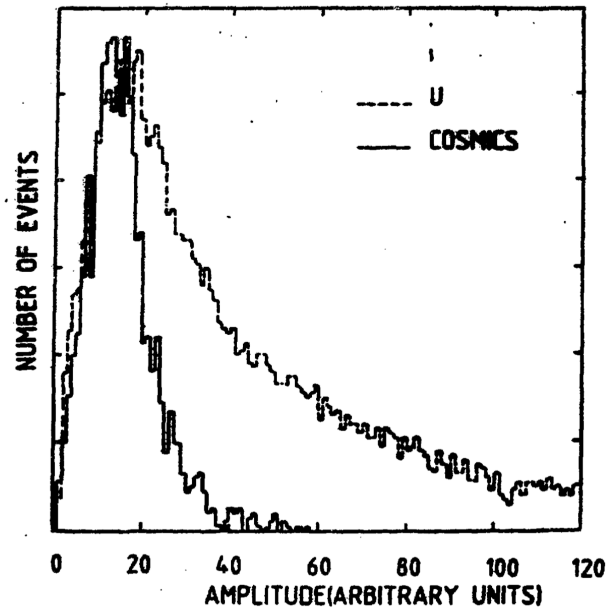






The IBM PC

- On August 12, 1981, IBM released their new computer, the IBM PC.
- In July of 1980, IBM representatives met with Microsoft's Bill Gates to talk about an operating system for the PC.
- In 1983, Time Magazine named the PC "Man of the Year."
  - i.e. just over 1 year from the launch.
- Jan 1986. IBM PC used for HEP L3 experiment at CERN to calibrate Hadron Calorimeter utilizing natural Uranium radioactivity (Laboratory of HEP, ITEP, Moscow)
- Now <at least> one laptop per physicist
  - ATLAS collaboration has 3000+ physicists and engineers
  - 2017 : Intel Core i7 ~320 000 MIPS

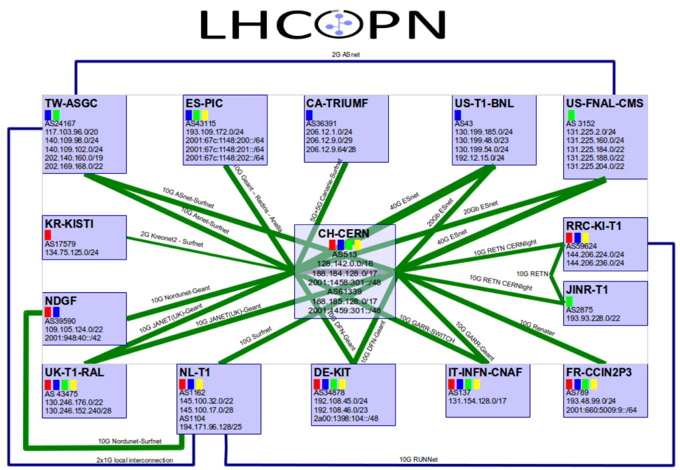


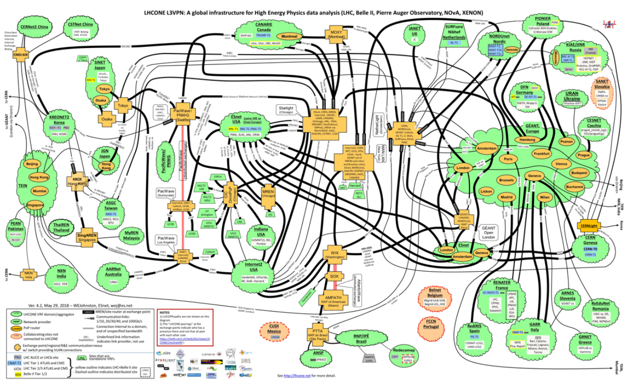
Act II

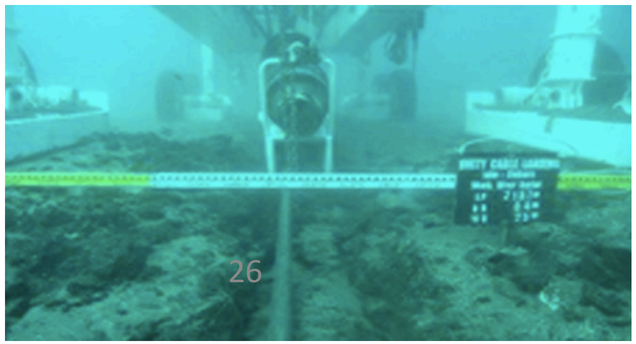
The Internet

Internet Timeline

- 1957: sputnik, ARPA
- Early 1960s: papers on packet switching, ideas for a “galactic network”
- Late 1960s: ARPANET
  - Original design speed: 2.4kbps
- Early 1970s: network control protocol
- 1 January 1983: move to TCP/IP
  - Originally 32 bit addresses
- 1986: US National Science Foundation develops NFSNET
  - Today leading backbone of scientific internet (ESNET)
- ~1999 LHC MONARC Model (to address Computing TDR’s concern not enough bandwidth
- ~2008 : LHC Optical Private Network to connect CERN and 12 main centers around the globe (10-40GB/s)
- ~2008 Google global network
- ~2011 LHC Optical network (LHCONE) to connect all (O(100)) centers

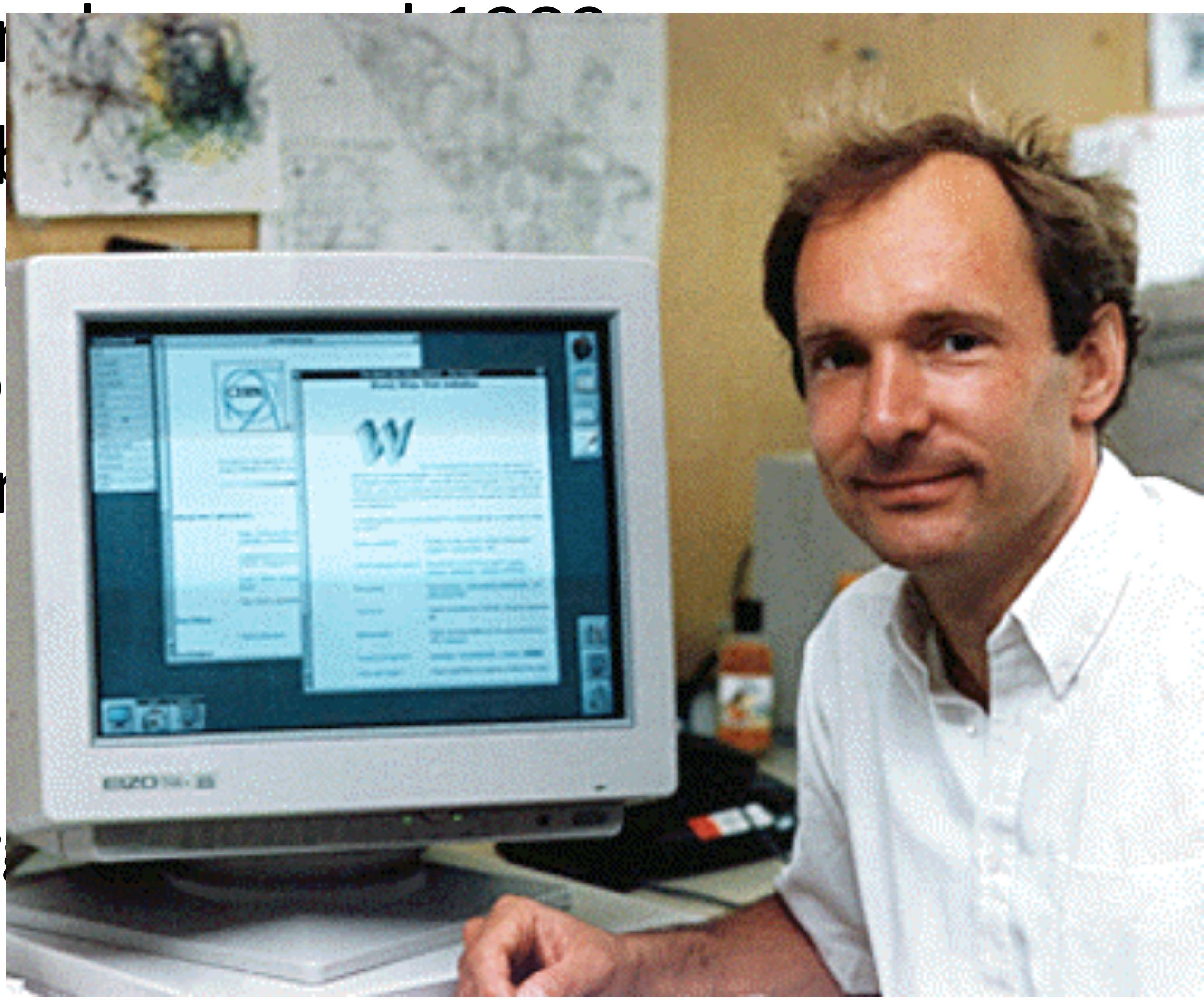






Birth of the Web

- Origi
- Resul
  'app
- Vario f
  Infor
- 1992
- 1993
  – Lar



WorldWideWeb : Proposal for a HyperText Project

WorldWideWeb: Proposal for a HyperText Project

To:

P.G. Innocenti/ECP, G. Kellner/ECP, D.O. Williams/CN

Cc:

R. Brun/CN, K. Gieselmann/ECP, R.€ Jones/ECP, T.€ Osborne/CN, P.

From:

T. Berners-Lee/CN, R. Cailliau/ECP

Date:

12 November 1990

First web proposal was on 12 March 1989
30 years ago !!!

The attached document describes in more detail a Hypertext project.

HyperText is a way to link and access information of various kinds as a web of nodes in which the user can browse at will. It provides a single user-interface to large classes of information (reports, notes, data-bases, computer documentation and on-line help). We propose a simple scheme incorporating servers already available at CERN.

The project has two phases: firstly we make use of existing software and hardware as well as implementing simple browsers for the user's workstations, based on an analysis of the requirements for information access needs by experiments. Secondly, we extend the application area by also allowing the users to add new material.

Phase one should take 3 months with the full manpower complement, phase two a further 3 months, but this phase is more open-ended, and a review of needs and wishes will be incorporated into it.

The manpower required is 4 software engineers and a programmer, (one of which could be a Fellow). Each person works on a specific part (eg. specific platform support).

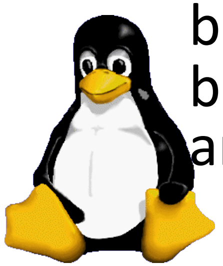
Each person will require a state-of-the-art workstation , but there must be one of each of the supported types. These will cost from 10 to 20k each, totalling 50k. In addition, we would like to use commercially available software as much as possible, and foresee an expense of 30k during development for one-user licences, visits to existing installations and consultancy.

We will assume that the project can rely on some computing support at no cost: development file space on existing development systems, installation and system manager support for daemon software.

T. Berners-Lee *R. Cailliau*

LINUX

- 1991 – the first version of Linux operating system
  - Posted announcement to the Minix group on USENET, and made the Linux source code available to other nerds free of charge. Programmers everywhere started adding their own improvements, and eventually companies like Red Hat, Corel, Caldera, and TurboLinux began selling their own versions of Linux.
  - The open-source nature of Linux is its greatest strength. Instead of having paid programmers devising improvements and looking for bugs from 9-to-5 with tight deadlines and budgets and memos from bosses, Linux is perpetually being tinkered with by the most obsessed and enthusiastic high-tech hobbyists and experts.



- 2011 – the primary Operating System in High Energy Physics Centers

'Just For Fun' is a humorous autobiography of Linus Torvalds

Larry Ellison (ORACLE CEO) predicted the future of computing (~early 2000):

- ***"There have been 3 generations of computing: mainframe, client-server and Internet computing***
- ***There'll be nothing new for one thousand (1000) years"***

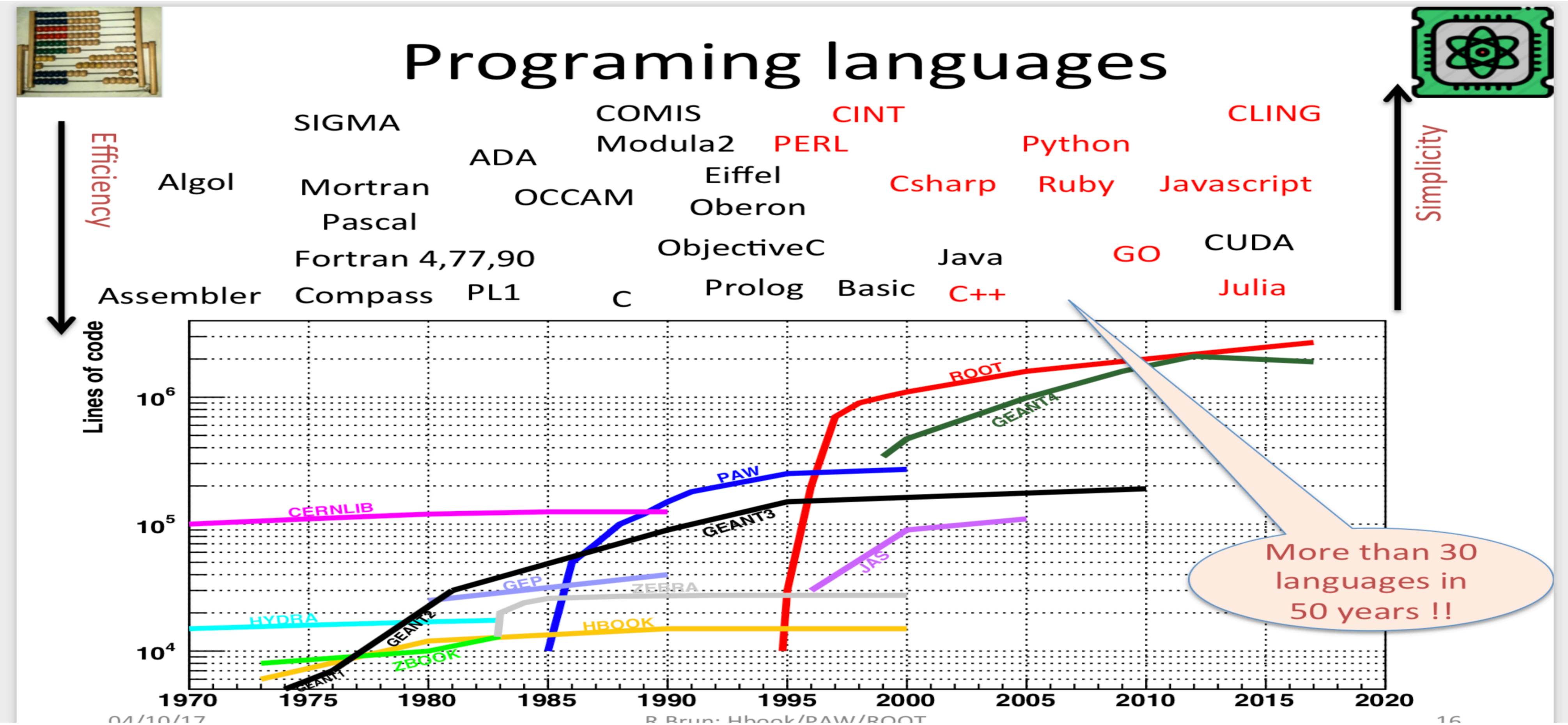
Curiously enough, very soon Oracle declared Grid to be **"the next big thing"**

Software Tools, Libraries, Languages

- 1974 : Fortran IV, HBOOK, GD3, HYDRA, CERNLIB, SIGMA
- 1979 : Fortran 77, Pascal, CERNLIB, GEANT2
- 1984 : ADA, ZEBRA, GEANT3, GHEISHA, PAW,X11
- 1989 : PHIGS, MOTIF, FORTRAN90, C, Perl, ORACLE, Web
- 1994 : OO, Smalltalk, Eiffel, C/C++, Objectivity, GEANT4, MOSAIC, ROOT
- 1999 : JAS, Mathematica, Netscape, PROOF, XROOTD
- 2004 : Google, Python
- 2014 : GPUs/AI, ML, DL, GO, CLING
- 2018 : python3
- 2020: C++20

What is next ? "Siri style" programming ?

An attempt to use commercial SW for LHC era



Этапы развития информационных технологий для экспериментов в области физики элементарных частиц

Годы	Число сотрудников эксперимента	Объем данных, технология хранения и обработки информации
Конец 1950	2-3	КБит'ы, записи в рабочих журналах
1960	10-15	КБайты, перфокарты, бумажные носители
1970 (У70, PS)	~35	МБайты, магнитные ленты Онлайн обработка : PDP 8, оффлайн обработка : ЕС, IBM 360
1980 (SPS, У70)	~100	Гигабайты, магнитные ленты и диски Онлайн обработка : Caviar, PDP 70, VAX, CM4, оффлайн обработка : ЕС, IBM 370, БЭСМ 6, VAX 8800
1990 (LEP, Теватрон, SLAC)	700-800	Терабайты, магнитные ленты, диски Онлайн обработка : VAX, спец.процессоры; Оффлайн обработка : ЕС, IBM 370, VAX 8800, Appollo, SGI, Sun
2010+ (LHC, AMS, Belle, LSST)	~3000	Петабайты, магнитные ленты, диски Онлайн обработка : кластеры, графические процессоры, Оффлайн обработка : грид
2030+ (HL-LHC, DUNE, EIC)	3000+	Эксабайты, магнитные ленты (?), диски, «облачные ресурсы» Онлайн обработка : специльные процессоры, кластеры Оффлайн обработка : «облачные ресурсы», грид, суперкомпьютеры

Timecheck

- 1969:
  - Man walks on the moon
  - End of first Star Trek series
  - First Arpanet nodes
- 1972:
  - First e-mail program
  - Telnet
  - Ethernet
- 1989:

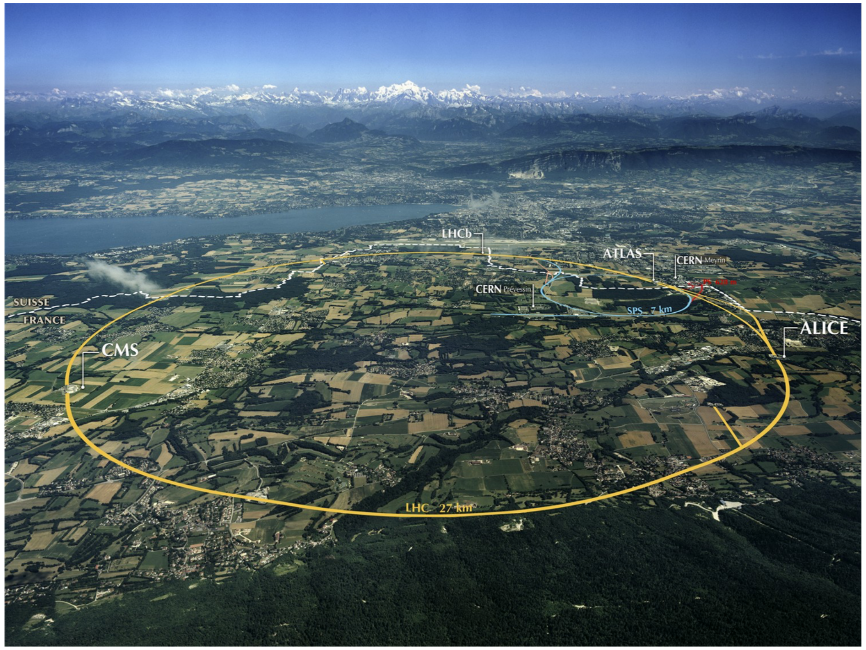
  \- Large Electron Positron Collider (LEP) at CERN pilot run

  LEP : 1989 - 2000
- 1991:
  - WWW protocols posted to alt.hypertext
- 1997:
  - Grid paradigm (by Foster and Kesselman)
- 1998:
  - The first "scheduled" LHC run
  - MONARC project
- 2000:
  - LEP stopped
- 2011:
  - LHC pilot run with Distributed Data processing at O(200) centers around the globe
- 2012-2013: LHC Physics Run (Run1). Higgs boson discovery
- 2015-2018 : LHC Physics Run (Run2)
  - **Exabytes of data processed and analyzed**
- 2020+ : HL-lHC challenges ahead

Act III

Computing at LHC, looking at the (Big)data

CERN

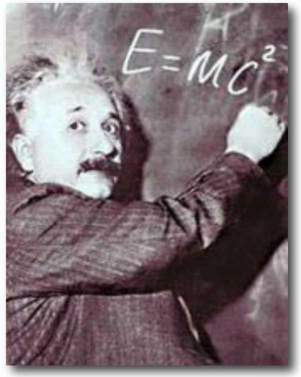


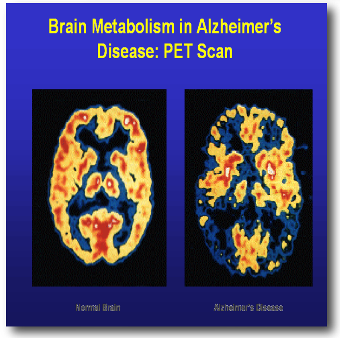
- **CERN is the European Organization for Nuclear Research in Geneva founded in 1954 : 12 European States**
  - **The world's largest particle physics laboratory**
  - **Particle accelerators and other infrastructure for high energy physics (HEP) and nuclear physics (NP) research**
  - **Worldwide community**
    - **23 member states (+ 2 members in pre-stage)**
    - **Observers: Russia, Japan, USA, EU, JINR and UNESCO**
    - **About 2700 staff, 20 experiments, 1200+ physicists**
    - **>10'000 users (about 5'000 on-site)**
    - **Budget (2019) ~1200 MCHF**

- **Birthplace of the World Wide Web**

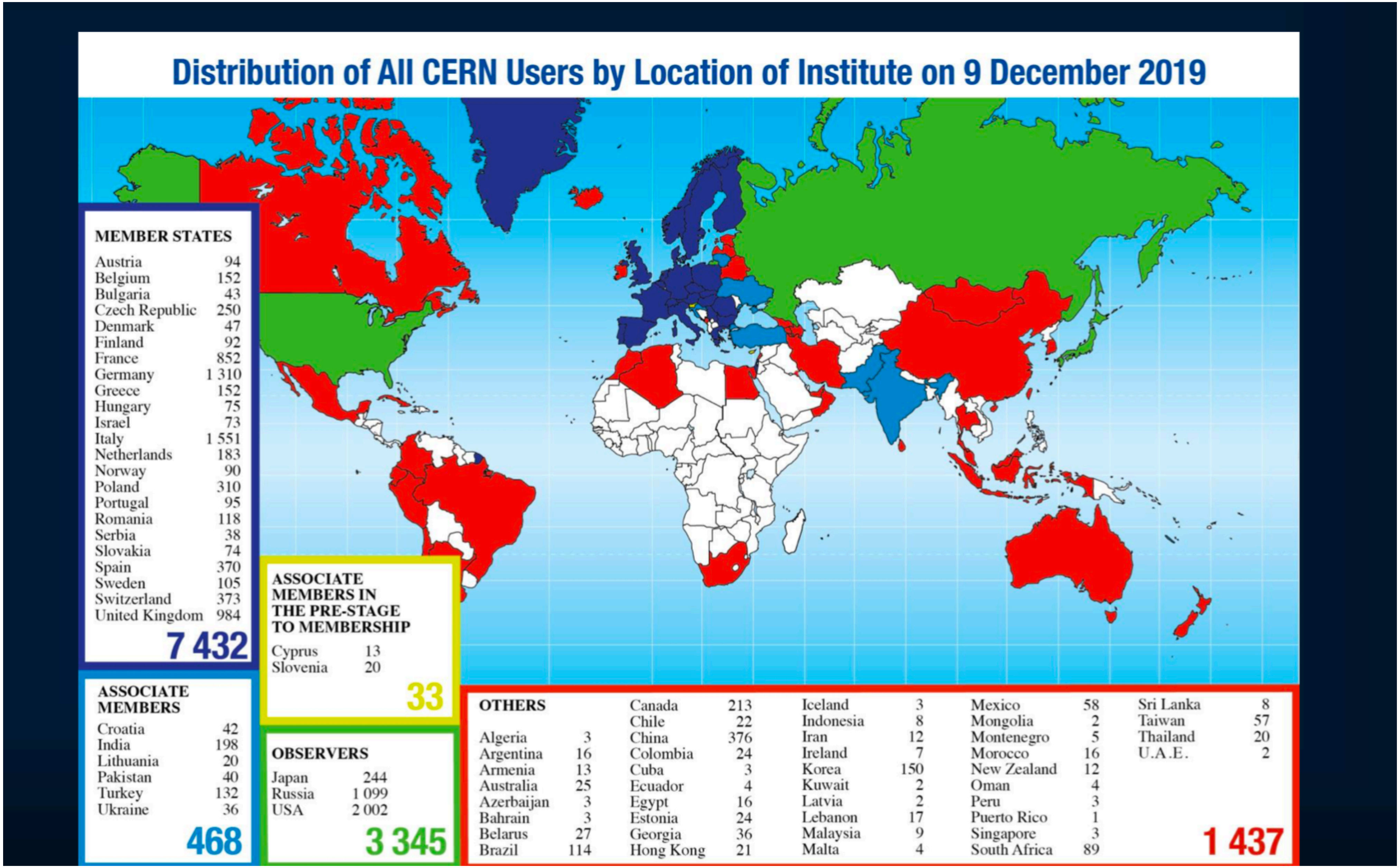
The Mission of CERN

- **Push forward the frontiers of knowledge**
  - The secrets of the Big Bang
  - Origin of mass
- **Develop new technologies for accelerators and detectors**
  - Information technology - the Web and the Grid
  - Medicine - diagnosis and therapy
- **Train scientists and engineers of tomorrow**
- **Unite people from different countries and cultures**

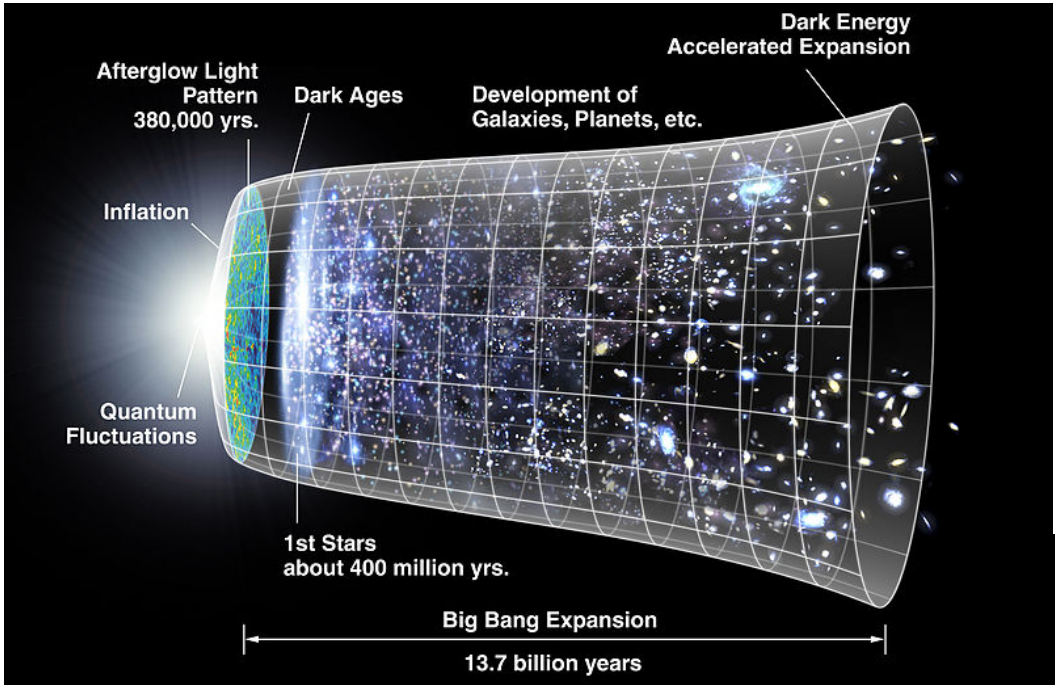




Science is getting more and more global

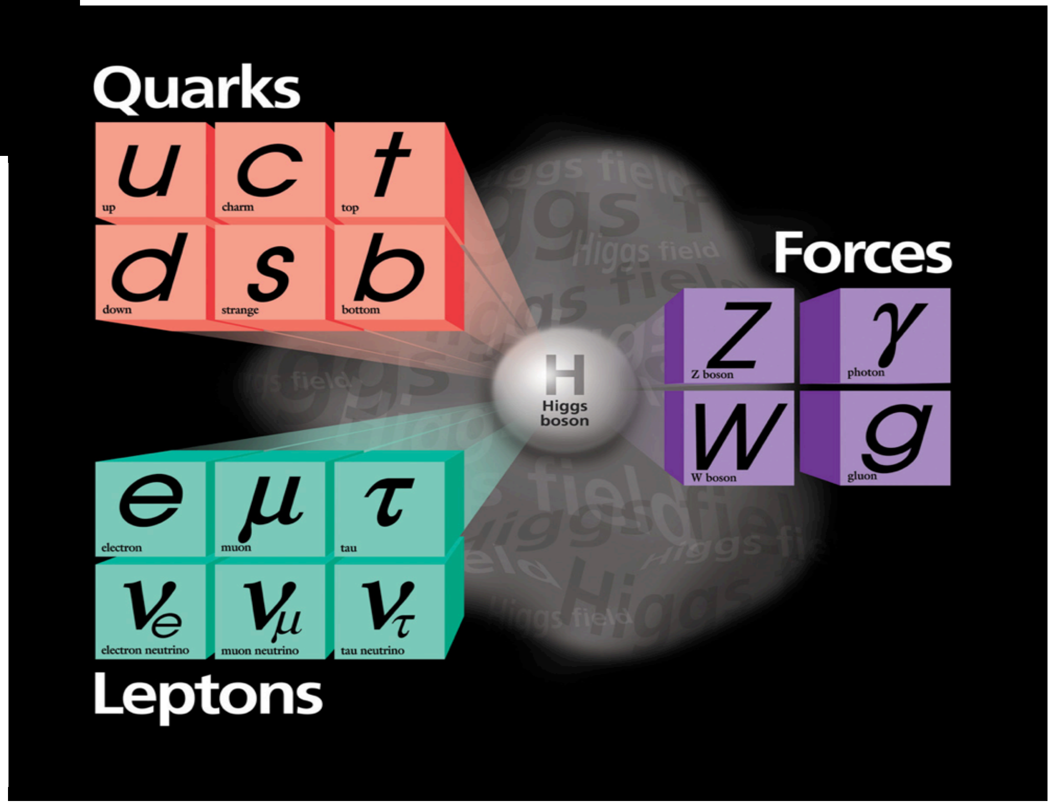


Fundamental Questions….



- How to explain that particles have mass?
  - Theories and accumulating experimental data…getting close
- What is 96% of the Universe made of?
  - We only observe 4% of the apparent mass

- Where is all the anti-matter?
  - Why is Nature not symmetric?
- What was the state of matter just after the Big Bang?
  - “Soup” of quarks and gluons before they condensed into matter?

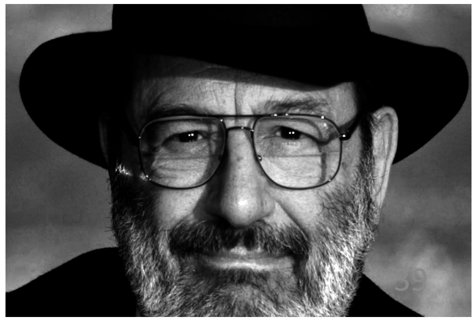


From Internet to Gutenberg

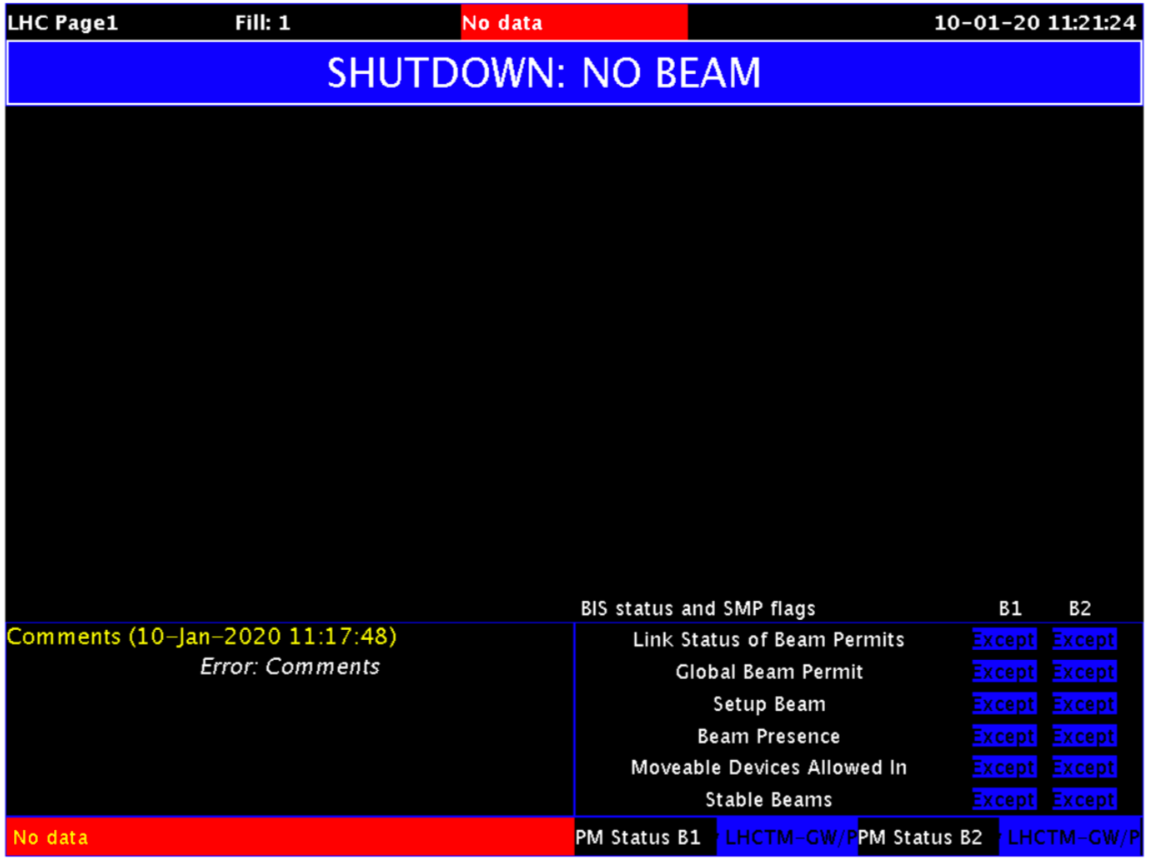
- Hermes, the alleged inventor of writing, presented his invention to the Pharaoh Thamus, he praised his new technique that was supposed to allow human beings to remember what they would otherwise forget.
- XX century (TV, radio, telephone…) brought another culture, people watch the whole world under the form of images which would have involved a decline of literacy.
- Computer screen is an ideal book on which one reads about the world in form of words and pages. If teen-agers by chance they want to program their own home computer, must know, or learn, logical procedures and algorithms, and must type words and numbers on a keyboard, at a great speed. In this sense one can say that the computer made us to return to a Gutenberg Galaxy.

- People who spend their night implementing an unending WWW conversation are principally dealing with words.

From lecture presented by Umberto Eco
(Italian philosopher, novelist,
author "Il nome della rosa")



What's happening now at CERN?



The Science Drivers for Particle Physics

Five intertwined science drivers, compelling lines of inquiry that show great promise for discovery :

1. *Use the Higgs boson as a new tool for discovery.*

2013

2. Pursue the physics associated with neutrino mass.

2015

3. *Identify the new physics of dark matter.*
4. Understand cosmic acceleration : dark energy and inflation.

2011

5. *Explore the unknown : new particles, interactions, and physical principles.*

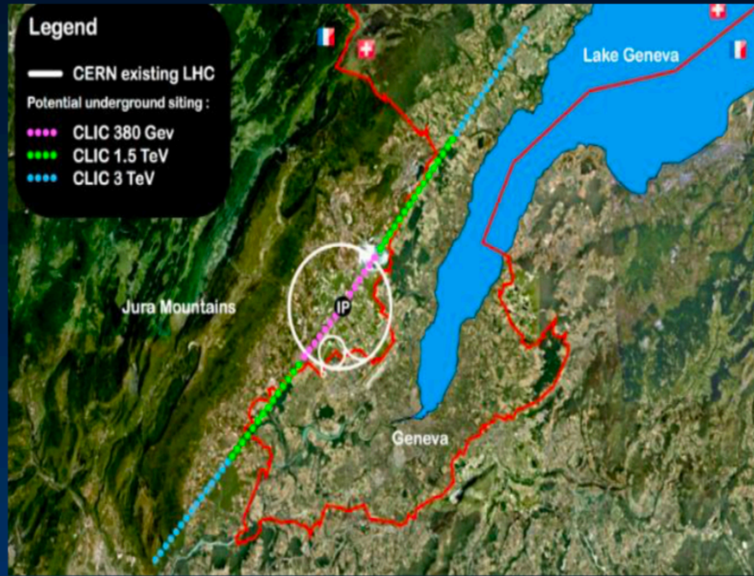
Future of particle physics

High Luminosity LHC until 2035

- Ten times more collisions than the original design

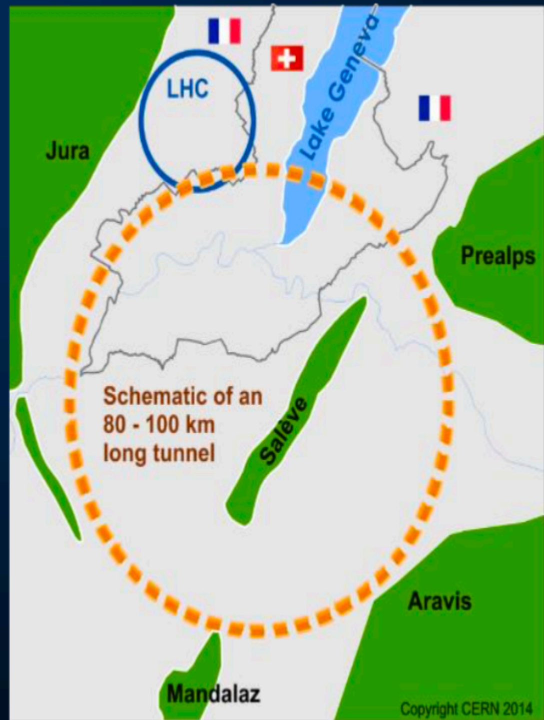
Studies in progress: Compact Linear Collider (CLIC)

- Linear e^+e^- collider $\sqrt{s}$ up to 3 TeV



Future Circular Collider (FCC)

- New technology magnets → 100 TeV pp collisions in 100km ring
- e^+e^- collider (FCC-ee) as 1st step?
- HE-LHC in the present LHC tunnel with FCC-hh technology?



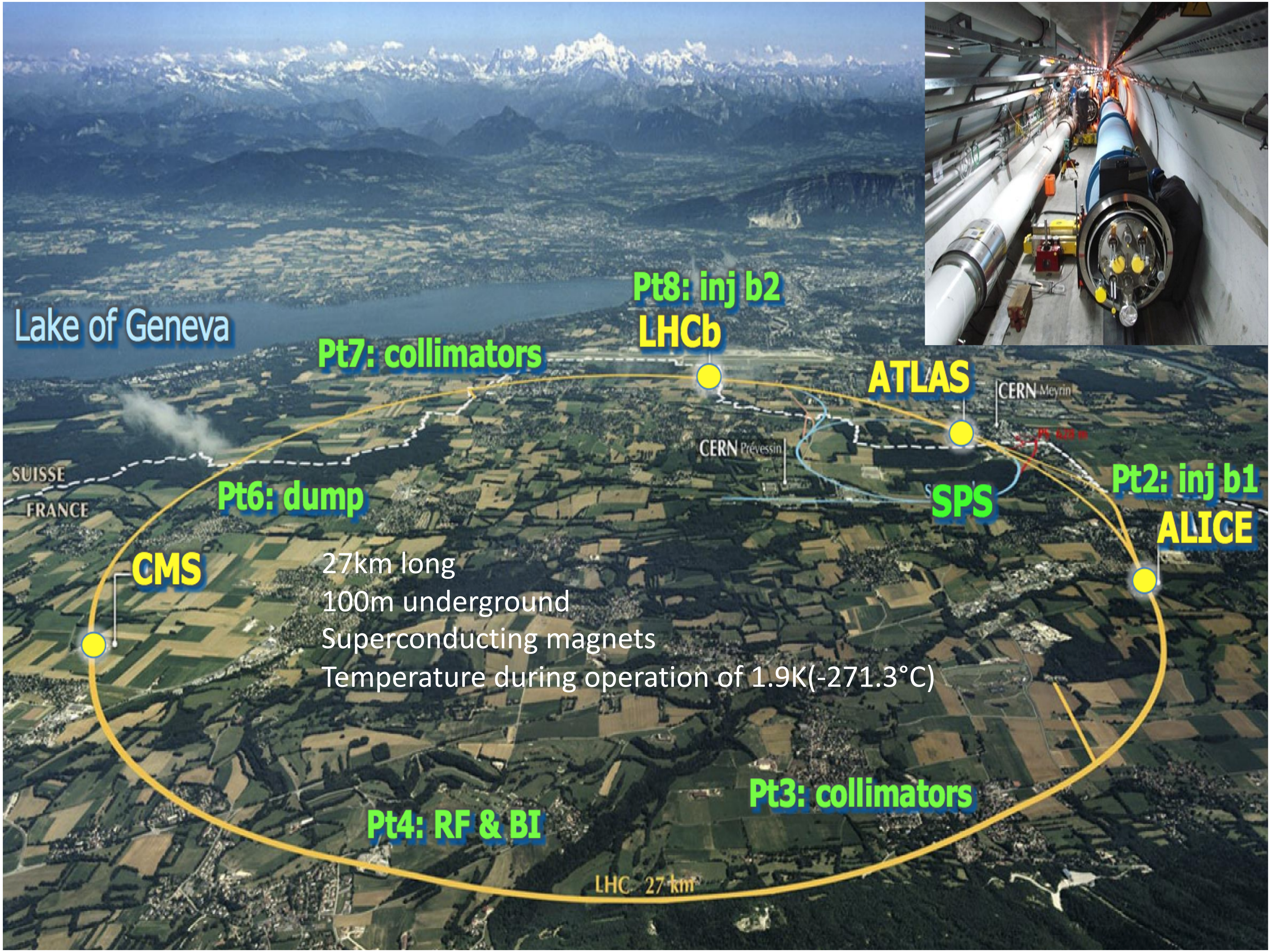
European Strategy for Particle Physics

- Preparing next update in 2020

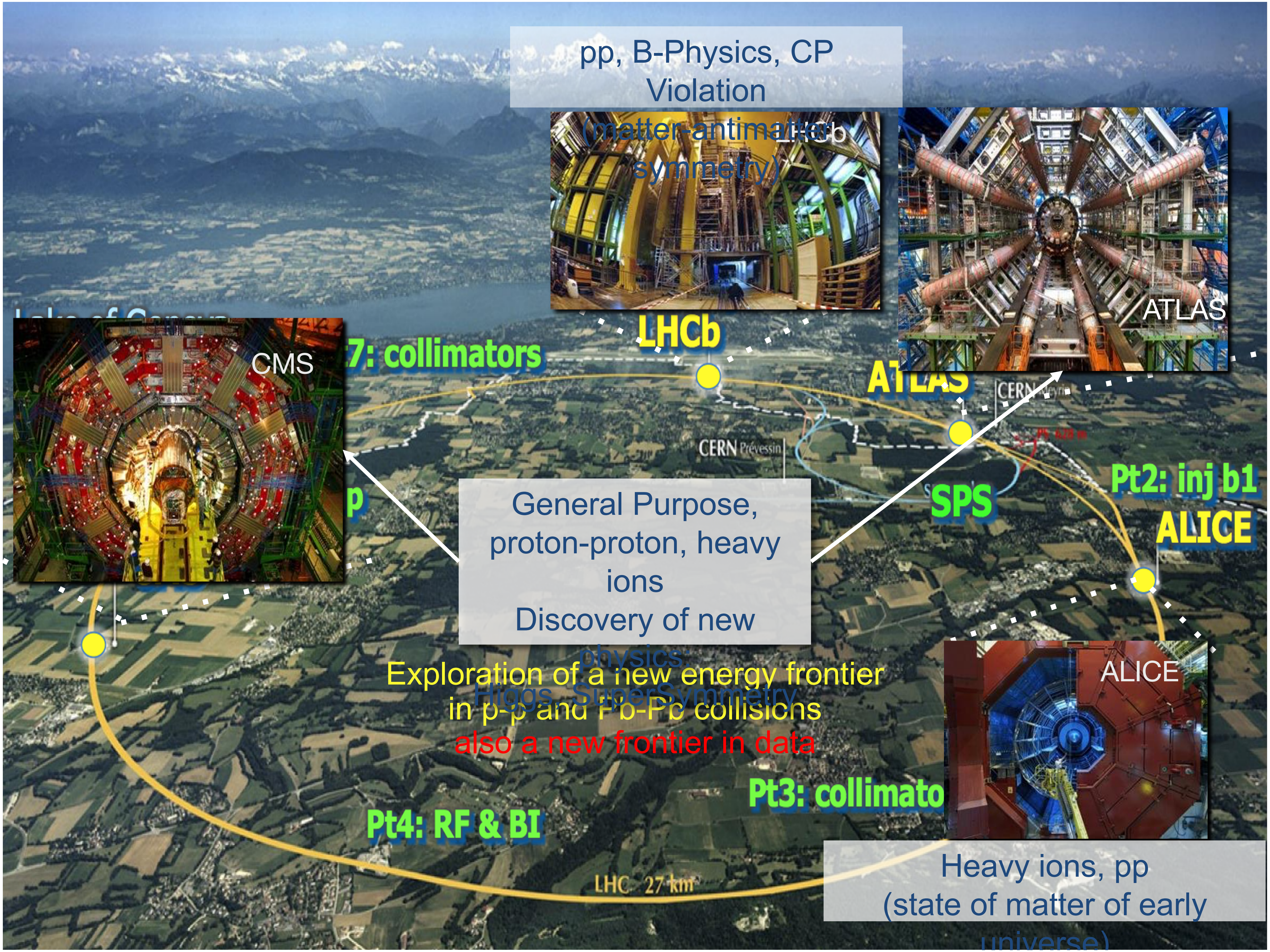
Prof. Sam Ting

History of Discoveries in Accelerators

Accelerator	Original purpose	Discovery
AGS Brookhaven (1960)	π N interactions	2 kinds of neutrinos, Breakdown of time reversal symmetry, 4-th Quark
FNAL Batavia (1970)	neutrino physics	5-th Quark, 6-th Quark
SLAC Spear (1970)	ep, QED	Partons, 4-th Quark, 3rd electron
PETRA Hamburg (1980)	6-th Quark	Gluons
Super Kamiokande (2000)	Proton Decay	Neutrino Oscillation
Hubble Space Telescope	Galactic Survey	Curvature of the universe
AMS on ISS	Dark Matter Antimatter	

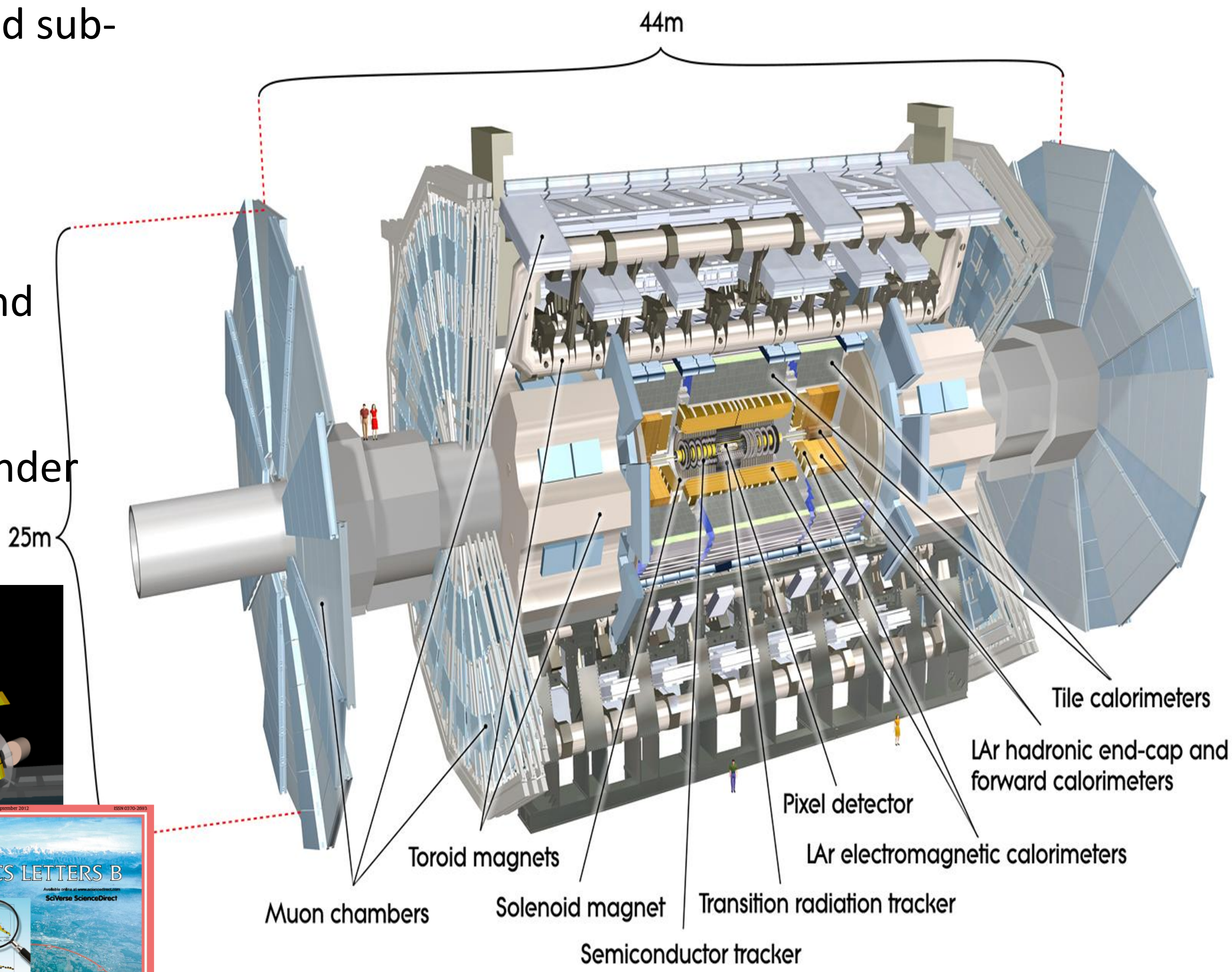


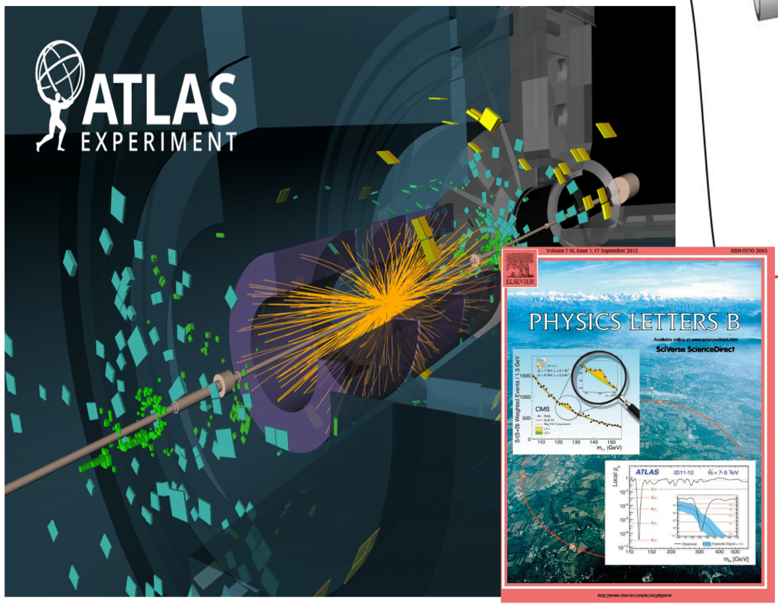
Lake of Geneva
Pt8: inj b2
LHCb
Pt7: collimators
ATLAS
CERN Meyrin
CERN Prévessin
SUISSE
FRANCE
Pt6: dump
SPS
Pt2: inj b1
ALICE
CMS
27km long
100m underground
Superconducting magnets
Temperature during operation of 1.9K(-271.3°C)
Pt3: collimators
Pt4: RF & BI
LHC 27 km


pp, B-Physics, CP
Violation
(matter-antimatter
symmetry)
LHCb
ATLAS
CMS
7: collimators
LHCb
ATLAS
CERN
CERN Prévessin
Pt2: inj b1
ALICE
SPS
General Purpose,
proton-proton, heavy
ions
Discovery of new
physics:
Higgs, SuperSymmetry
Exploration of a new energy frontier
in p-p and Pb-Pb collisions
also a new frontier in data
ALICE
Pt3: collimato
Pt4: RF & BI
LHC 27 km
Heavy ions, pp
(state of matter of early
universe)

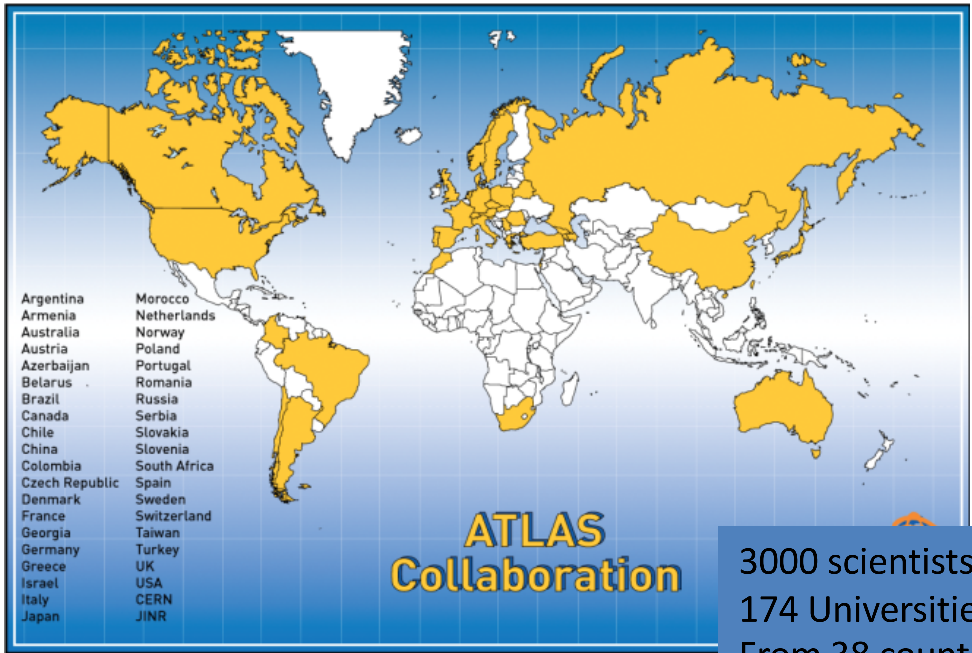
ATLAS

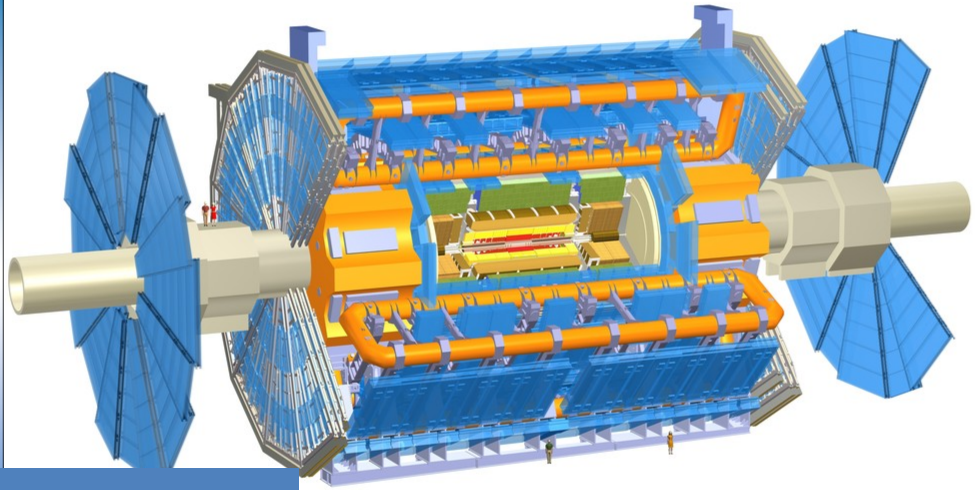
- The largest detector
- Multiple components and sub-detectors
- 7000 tons
- 10 MW electric power
- 150M sensors measure direction, momentum and charge
- Collisions at 40MHz
    - Filtered to kHz or under 6GB/s





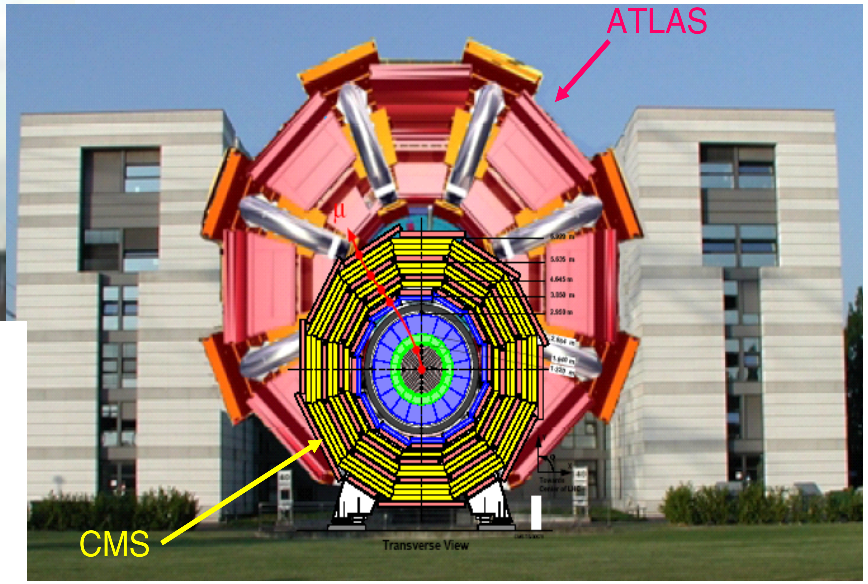
The ATLAS Experiment at the LHC

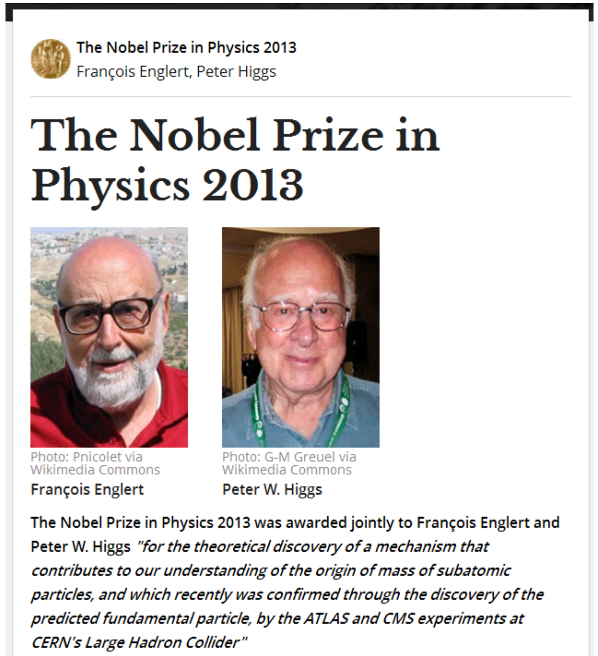




3000 scientists
174 Universities and Labs
From 38 countries
More than 1200 students

- **ATLAS has 44 meters long and 25 meters in diameter, weighs about 7,000 tons. It is about half as big as the Notre Dame Cathedral in Paris and weighs the same as the Eiffel Tower or a hundred 747 jets**





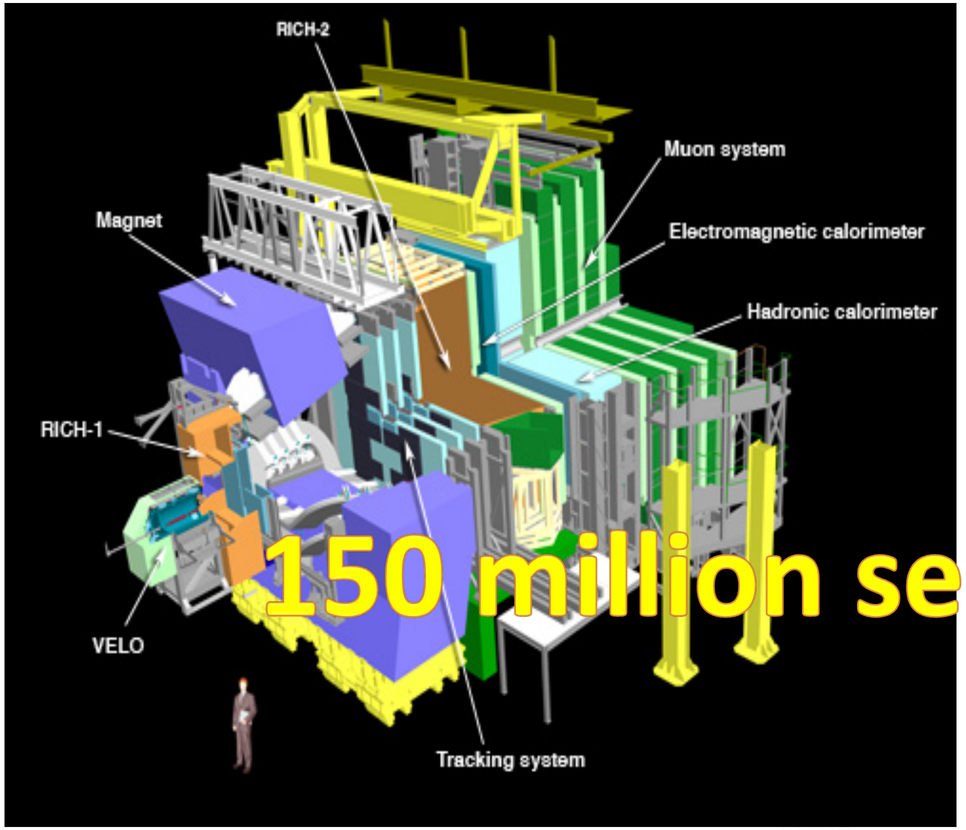
The Nobel Prize in Physics 2013
François Englert, Peter Higgs

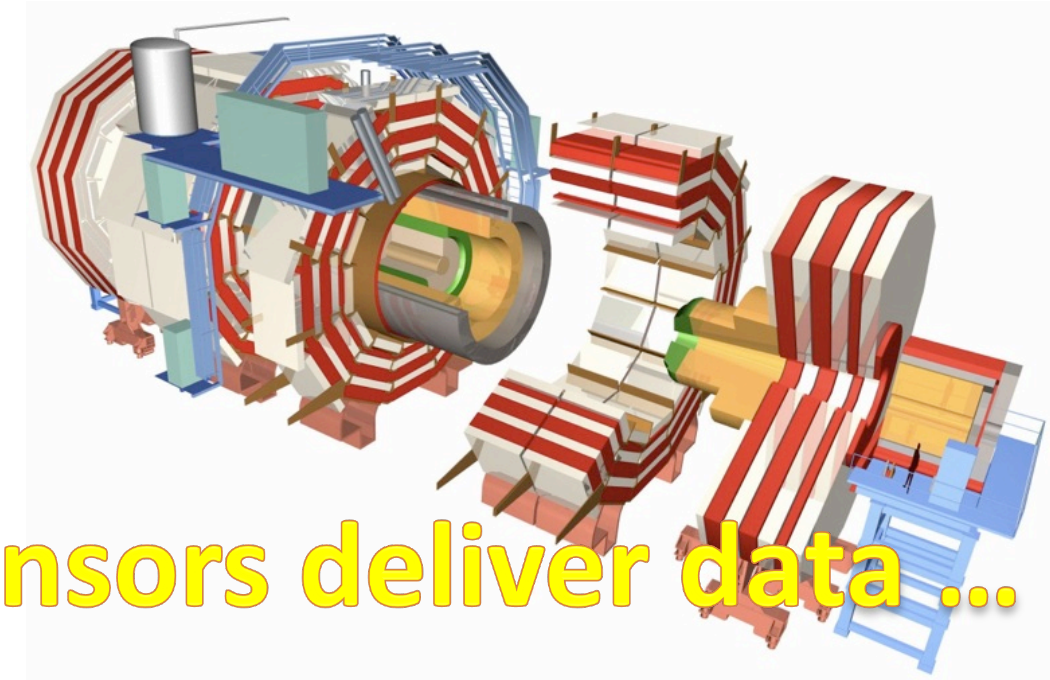
The Nobel Prize in Physics 2013

Photo: Pnicolet via Wikimedia Commons
François Englert

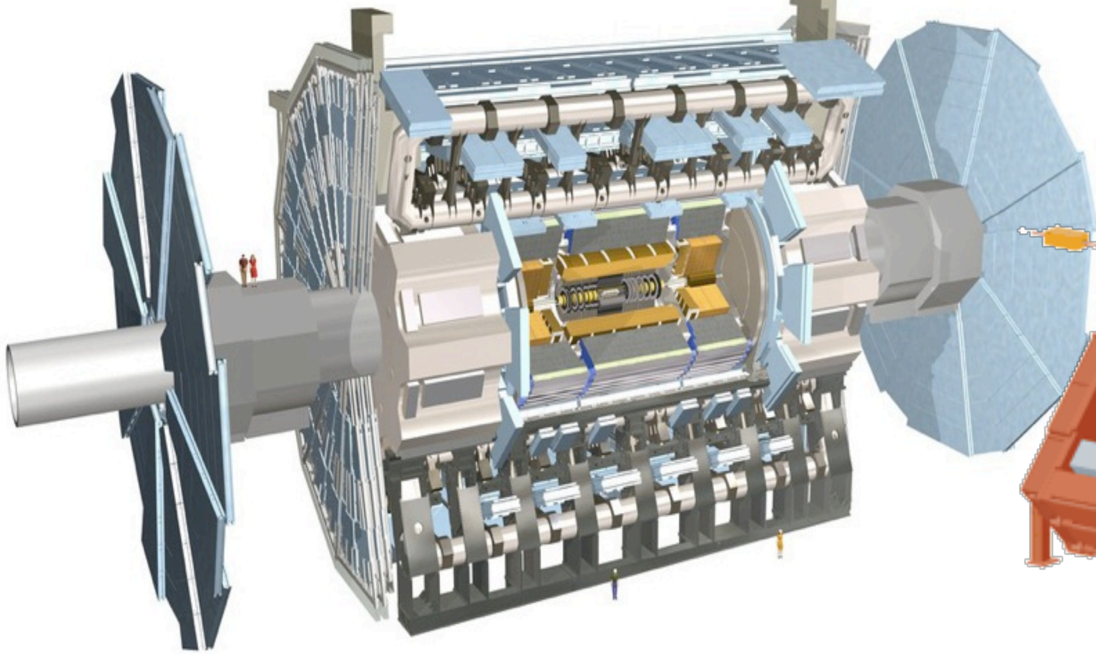
Photo: G-M Greuel via Wikimedia Commons
Peter W. Higgs

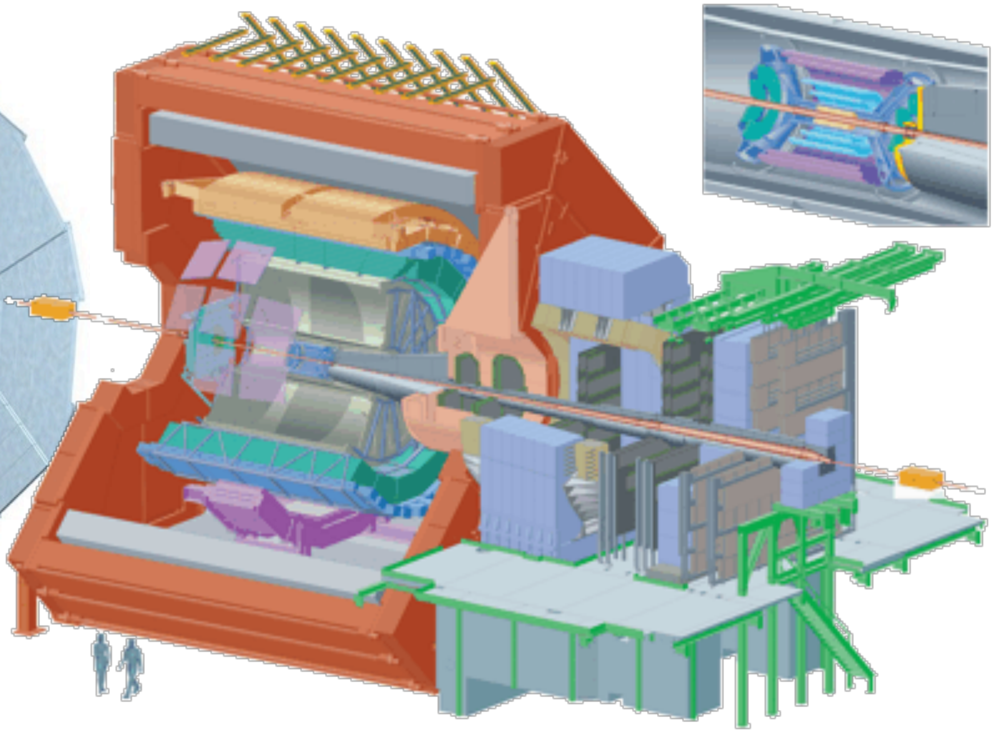
The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*



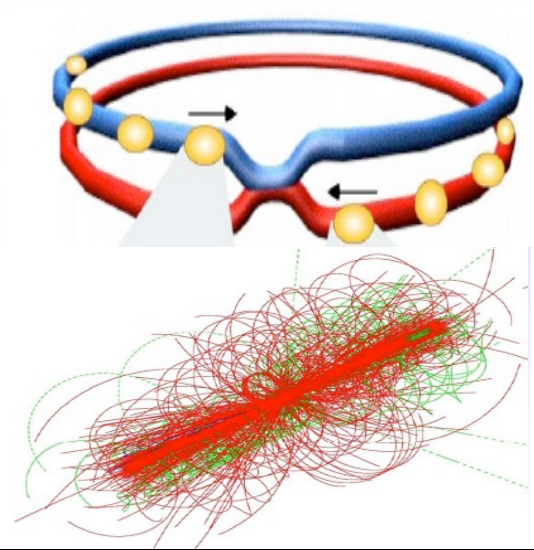


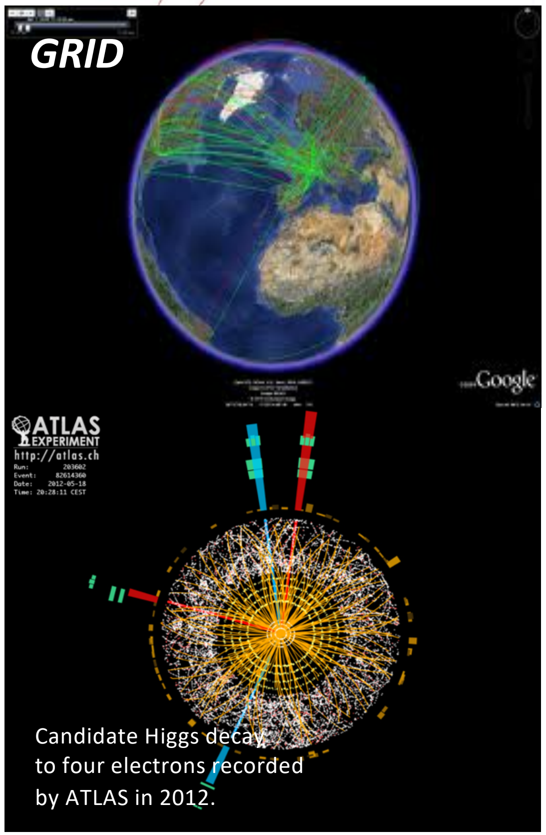
150 million sensors deliver data ...



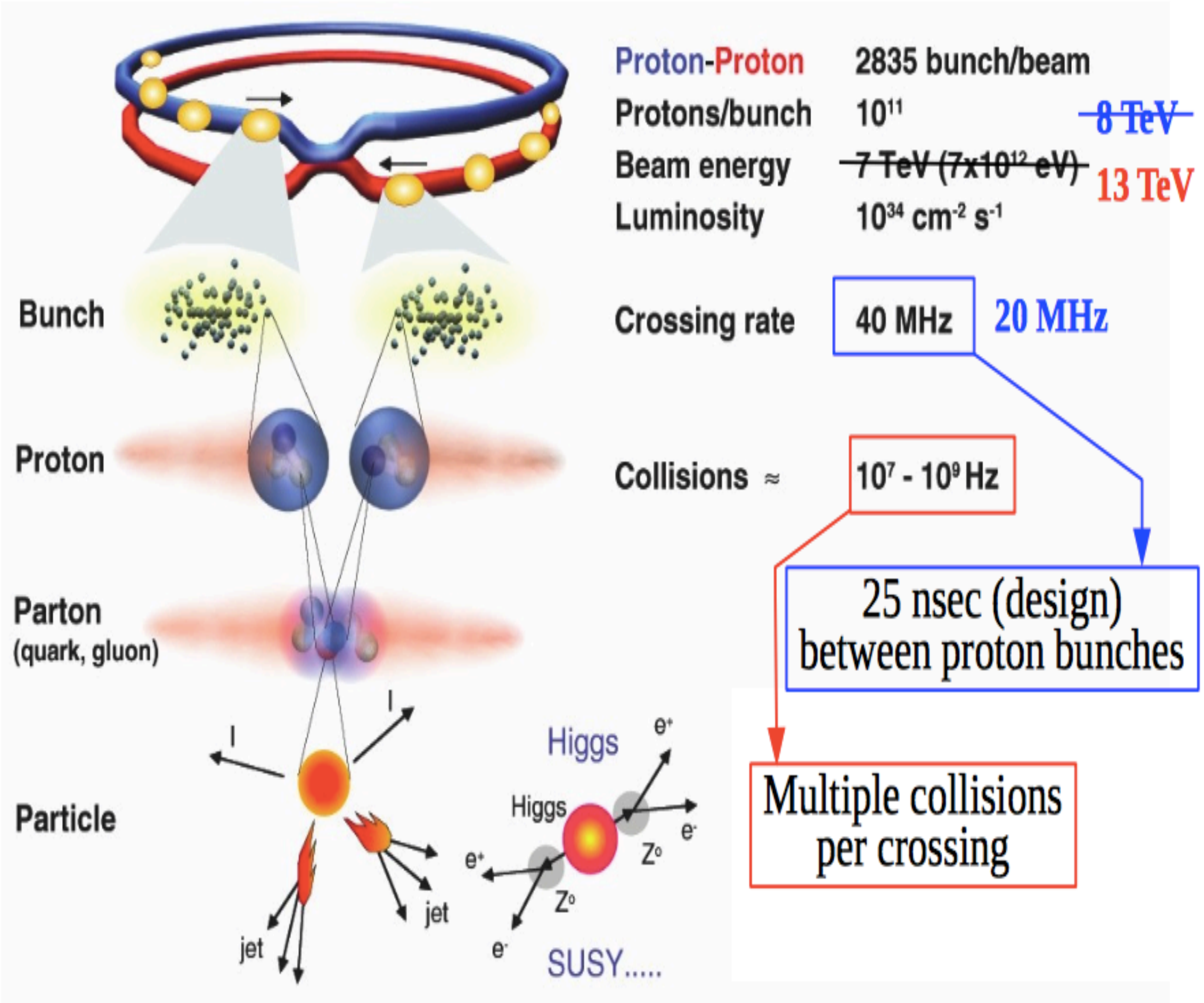


Proton-Proton Collisions at the LHC





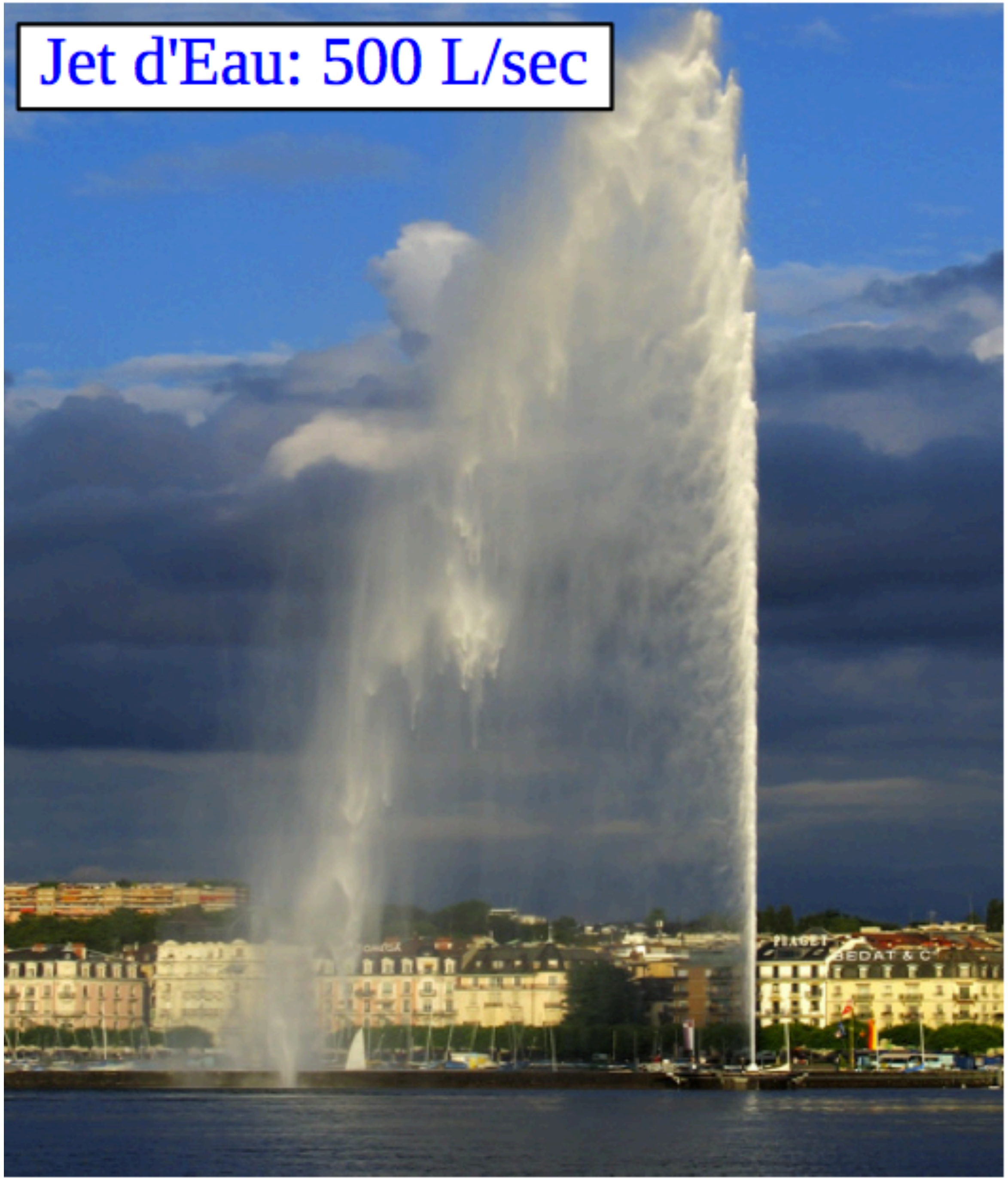
A Few LHC Facts



New physics rate ~ 0.00001 Hz

Event Selection : 1 in 10,000,000,000,000

New physics rate ~ 0.00001 Hz



Event Selection :

1 in 10,000,000,000,000

Like looking for a single drop of water from the Geneve Jet d'Eau over 2+ days

Information growth

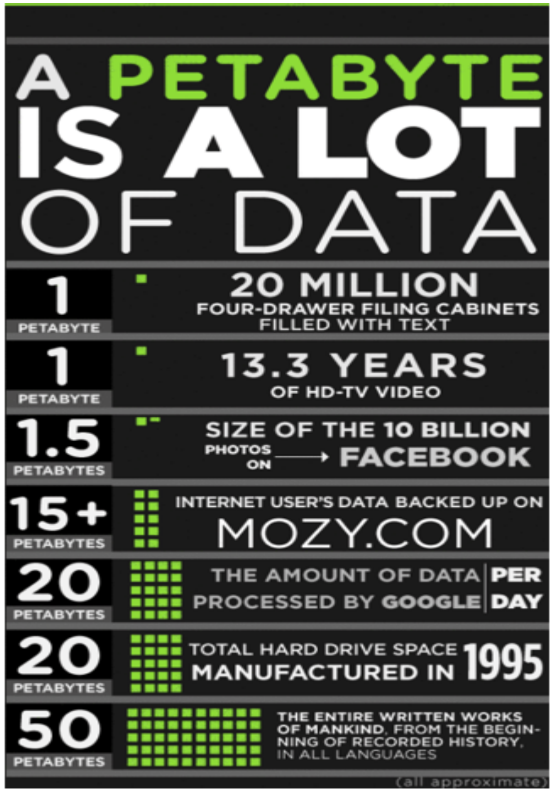
Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

Library of congress ~ 200 TB

One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
30 Petabyte * N for copies

One LHC event ~ 2 MByte yearly → 10 billion RAW events
Tycho's rate ~100 Byte/h ~ 800 Petabytes with copies, derived data and selections

One photo ~ 2 Mbytes yearly → 500 billion photos
one Exabyte
25 billion photos on Facebook

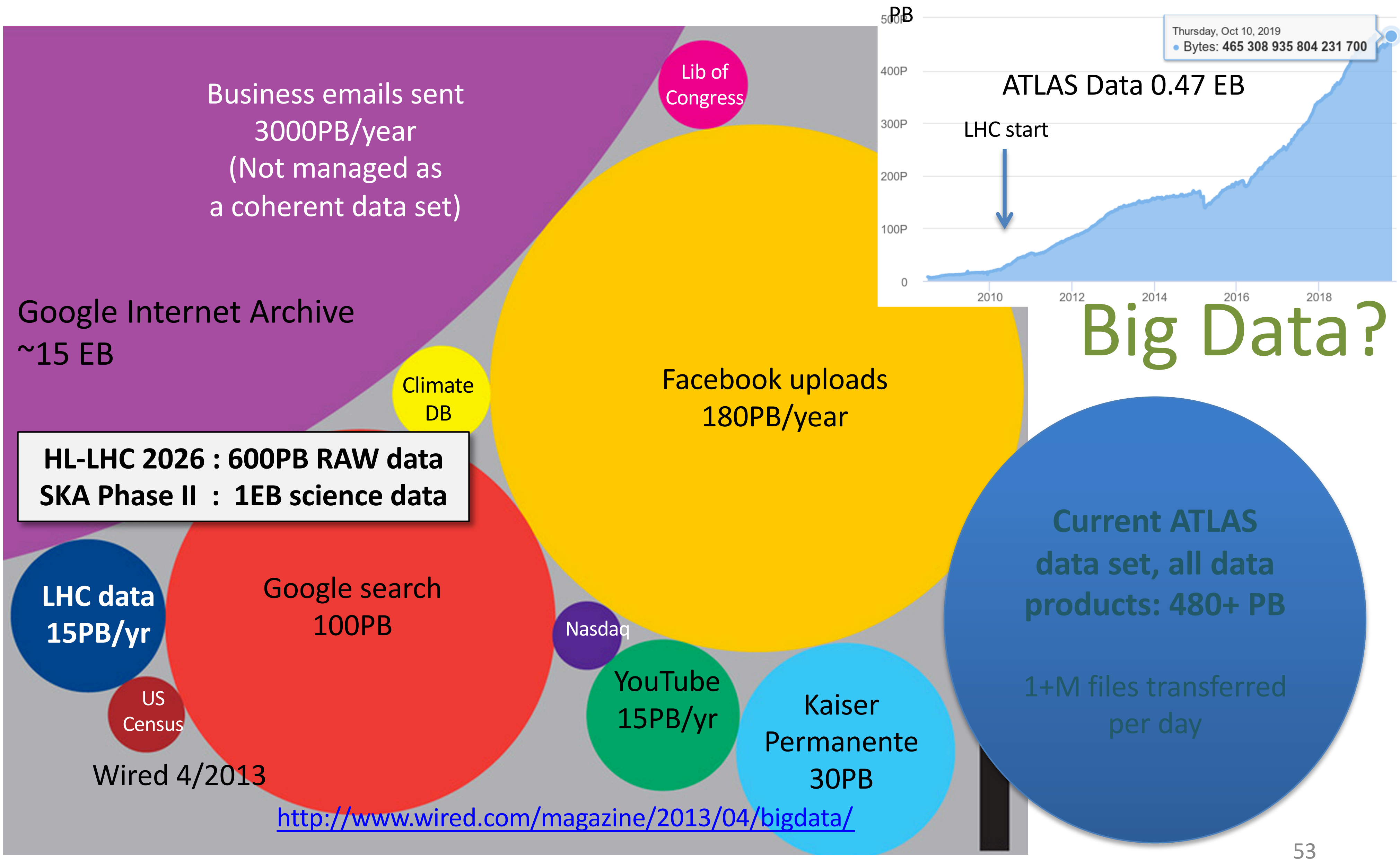


World Wide Web 25 billion pages (searchable)
Deep web estimates > 1 trillion documents ~ one Exabyte

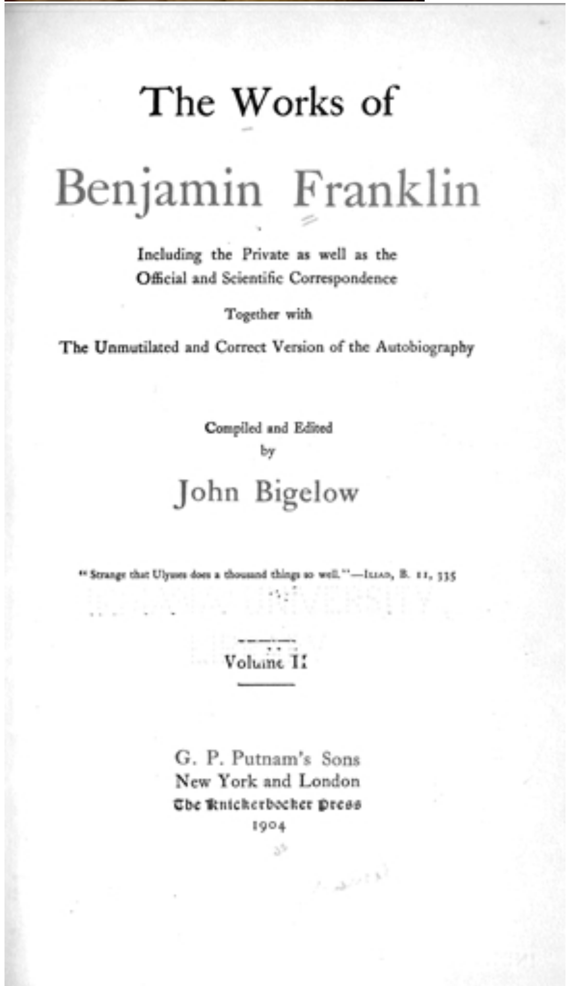
Blue-ray disks ~ 25 Gbytes yearly → 100 million units
2.5 Exabytes, copies

World-wide telephone calls → ~ 50 Exabyte (ECHELON)

Big Data: often just a buzz word, but not when it comes to HEP…



Big Data in the arts and humanities



The Works of

Benjamin Franklin

Including the Private as well as the
Official and Scientific Correspondence

Together with

The Unmutilated and Correct Version of the Autobiography

Compiled and Edited
by

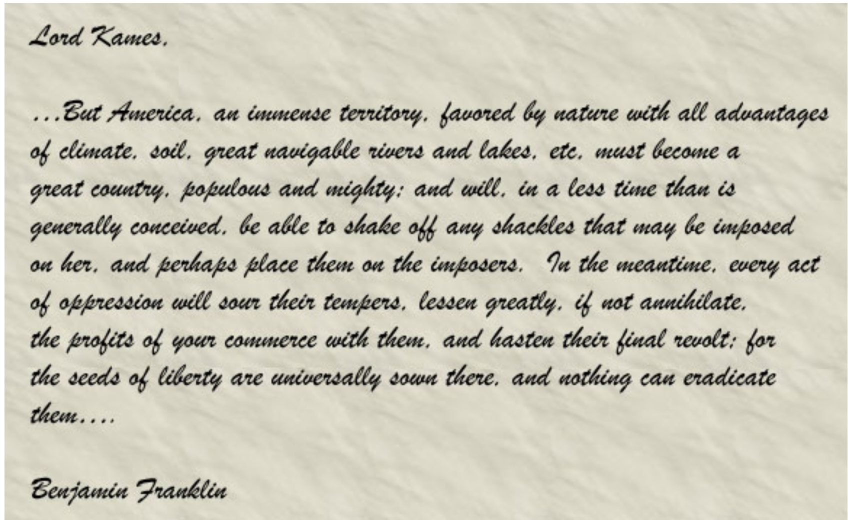
John Bigelow

"Strange that Ulysses does a thousand things so well."—Iliad, B. ii, 335

Volume II

G. P. Putnam's Sons
New York and London
The Knickerbocker Press
1904

Letter of Benjamin Franklin to Lord Kames, April 11, 1767. Franklin warned British official what would happen if the English kept trying to control the colonists by imposing taxes, like the Stamp Act. He warned that they would revolt.



Lord Kames,

...But America, an immense territory, favored by nature with all advantages of climate, soil, great navigable rivers and lakes, etc. must become a great country, populous and mighty; and will, in a less time than is generally conceived, be able to shake off any shackles that may be imposed on her, and perhaps place them on the imposers. In the meantime, every act of oppression will sour their tempers, lessen greatly, if not annihilate, the profits of your commerce with them, and hasten their final revolt; for the seeds of liberty are universally sown there, and nothing can eradicate them....

Benjamin Franklin

The political, scientific and literary papers of Franklin comprise approx. 40 volumes containing approx. 100,000 documents.

Big Data in the arts and humanities

George W. Bush Presidential Library:
200 million e-mails
4 million photographs

Терминология

Компьютерная модель *(вычислительная модель, модель компьютинга)* : модель обработки данных физического эксперимента.

Сверхбольшие объемы данных : данные мультипетабайтного и эксабайтного диапазона в современных физических экспериментах. В отличии от «больших данных» - этот термин в настоящее время имеет специальное значение и подразумевает использование целого стека соответствующих технологий.

Популярность данных *(от англ. data popularity)* – востребованность данных и их наборов для программ физического анализа или изучения работы детектора.

«облако ATLAS» – статическая связка центров уровня 1:T1 – n:T2 в рамках иерархической компьютерной модели на первом этапе работы Большого адронного коллайдера

Terminology. Cont'd

- Data is collected online (real-time)
  - Collision data recorded by the detectors
- Physicists analyze this data offline
  - Optimizing selection, estimating/modeling background, establishing limits, discovering New Physics, etc.
- The LHC delivers a lot of data, which we need first to select online for future analysis
  - Data filtering is done online and offline

Роль информационных технологий в физике элементарных частиц

Исследования в области физики высоких энергий (ФВЭ) и ядерной физики (ЯФ) невозможны без использования значительных вычислительных мощностей и программного обеспечения для обработки, моделирования и анализа данных.

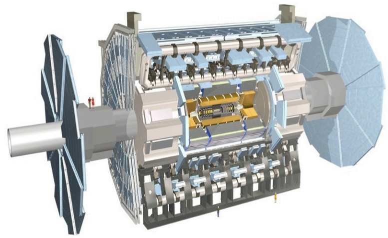
Это определяется рядом факторов:

- большими объемами информации, получаемыми с установок на современных ускорителях;
- сложностью алгоритмов обработки данных (ATLAS SW 5M LoC);
- статистической природой анализа данных;
- необходимостью (пере)обрабатывать данные после уточнения условий работы детекторов и ускорителя и/или проведения калибровки каналов считывания;
- необходимостью моделирования условий работы современных установок и физических процессов одновременно с набором и обработкой «реальных» данных.

Введение в строй Большого адронного коллайдера, запуск установок масштаба ATLAS, CMS, эксперименты на будущих установках, таких как NICA, характеризуются сверхбольшими объемами информации и тысячами ученых ведущих исследования. Это требует новых подходов, методов и решений в области информационных технологий.

Physics Dataflow

One Experiment



Detector →
150 million electronic channels

1 PBytes/s

Level 1 Filter and Selection

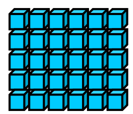
Fast response electronics, FPGA, embedded processors, very close to the detector

Limits :
Essentially the budget and the downstream data flow pressure

150 GBytes/s

High Level Filter and Selection

O(1000) PC's for processing, Gbit Ethernet Network

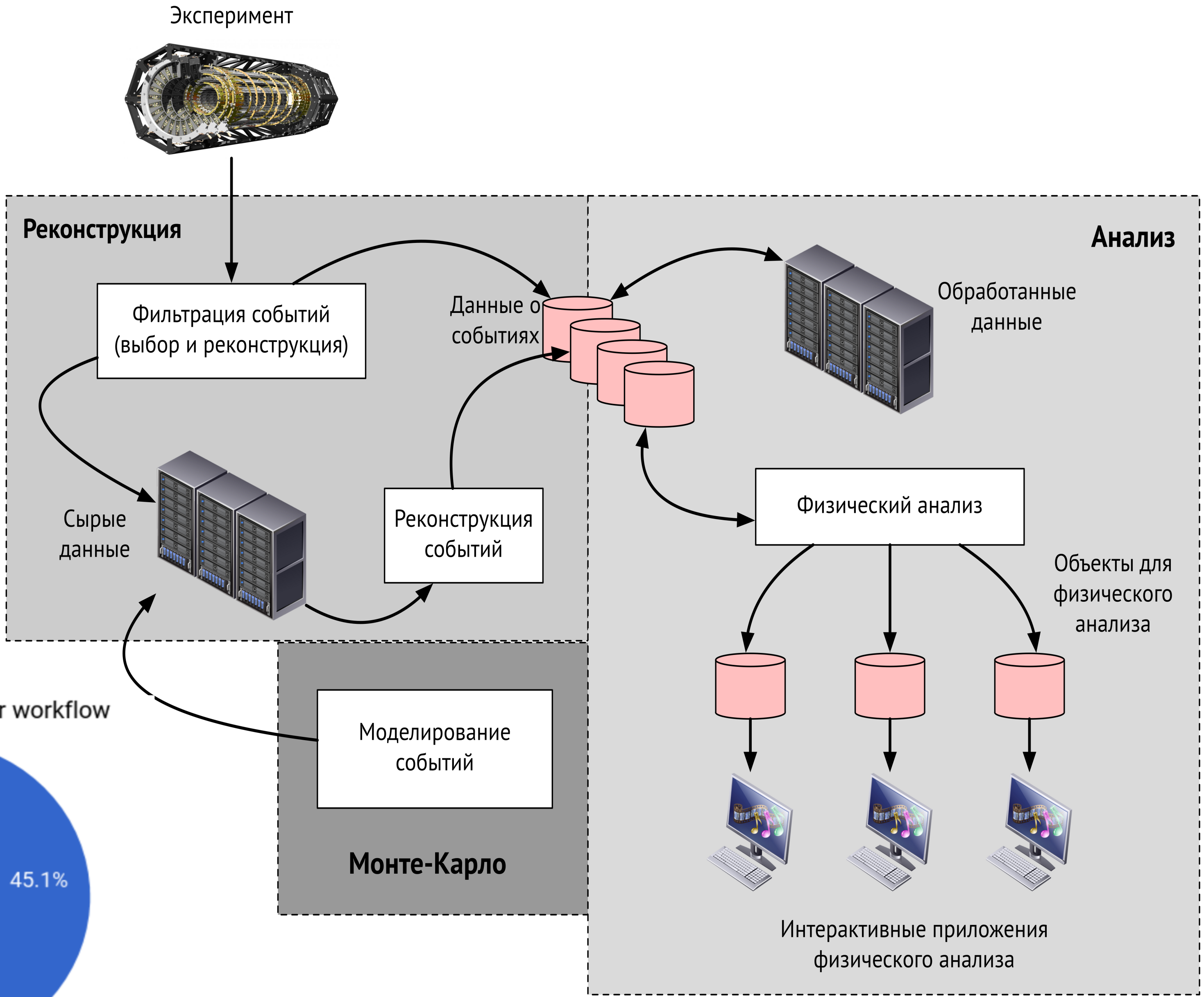


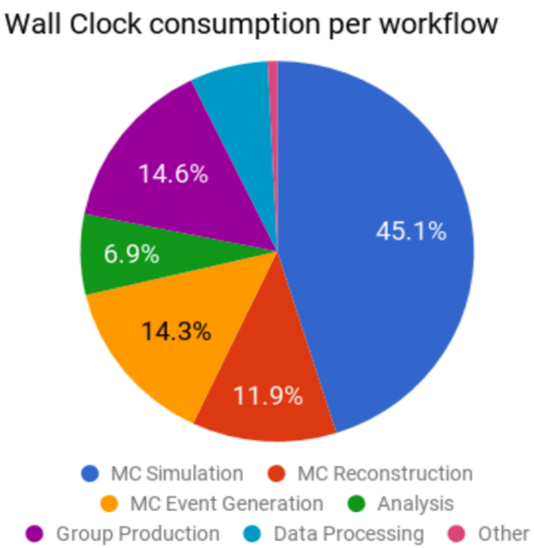
N x 10 Gbit links to the Computer Center

0.6 GBytes/s

CERN Computer Center

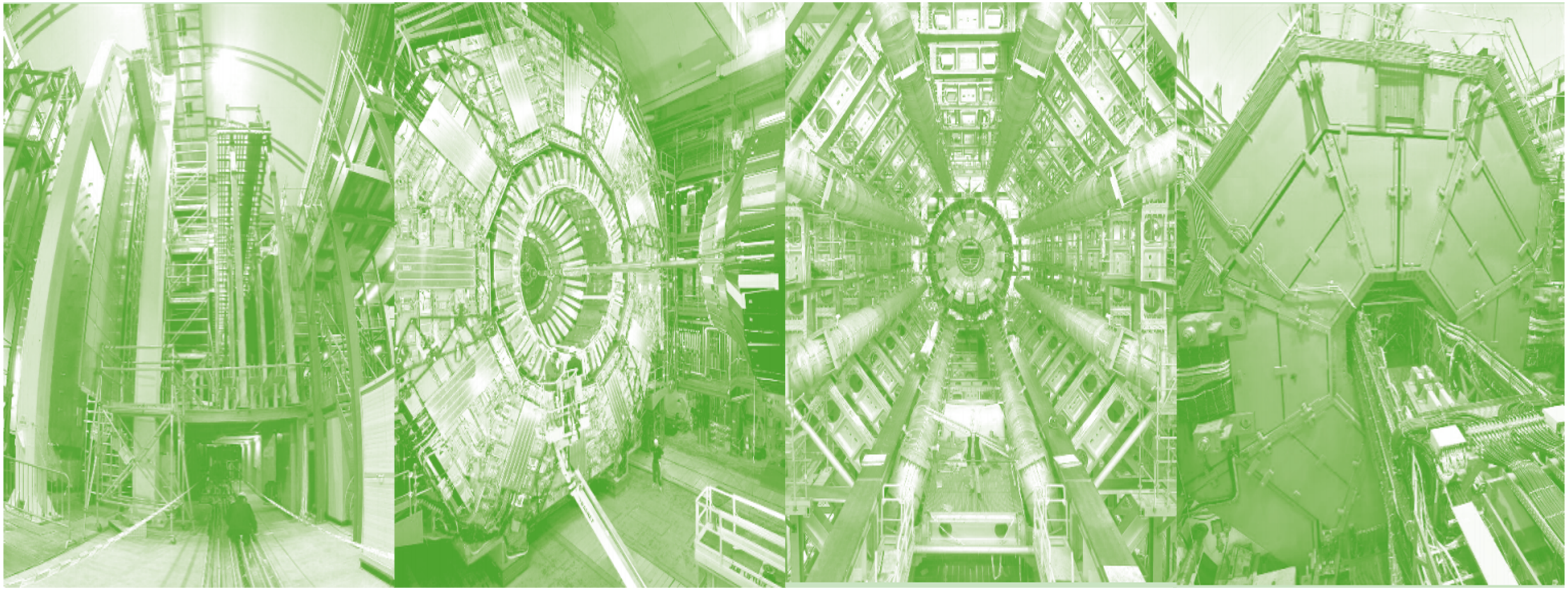
Роль информационных технологий в Физике Высоких Энергий и Ядерной Физике



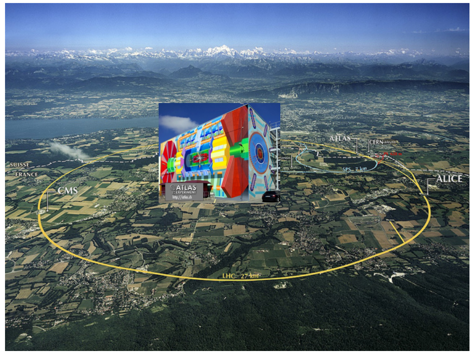


Реальное время, затрачиваемое на выполнение задач различных типов

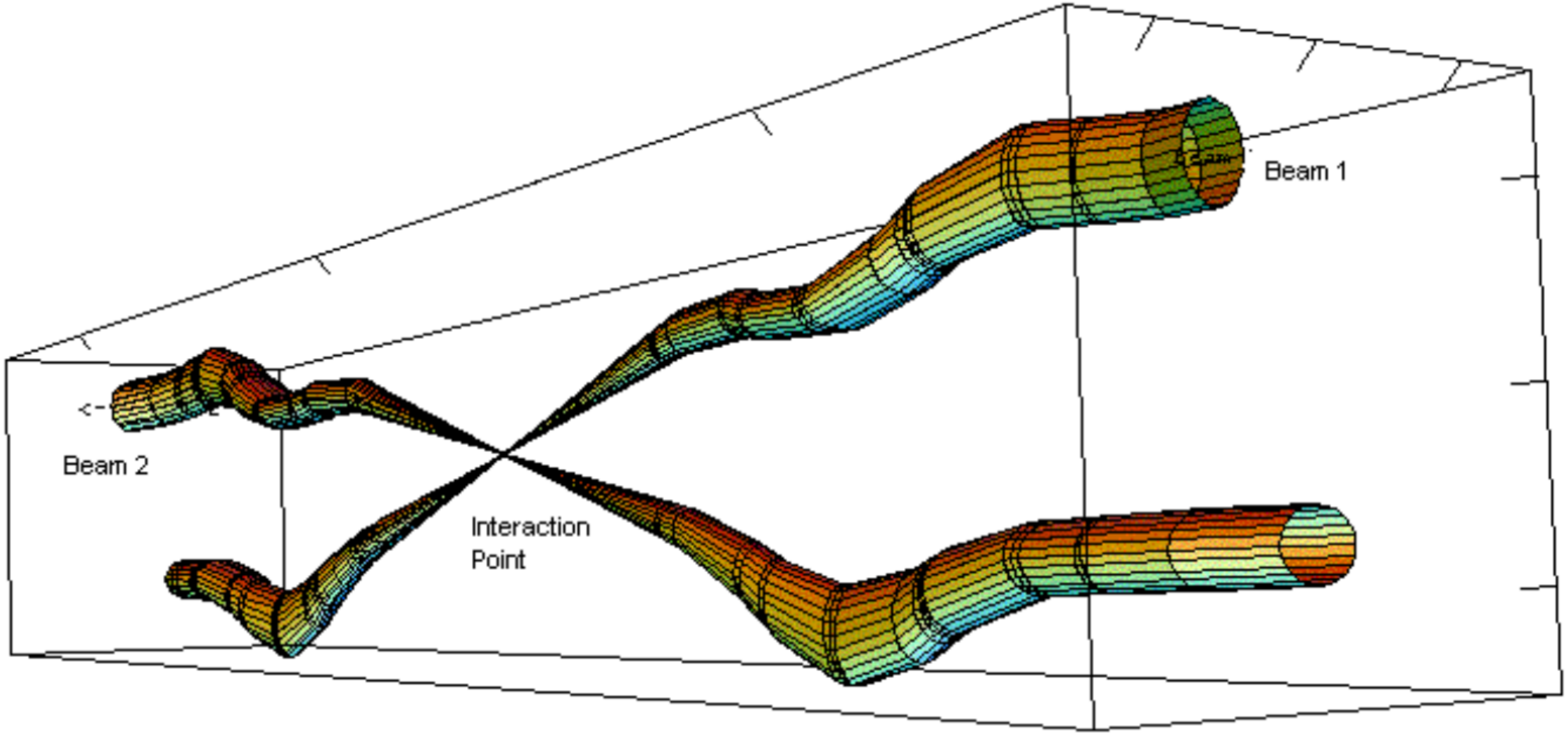
LHC Challenges



WHAT IS AN EVENT?

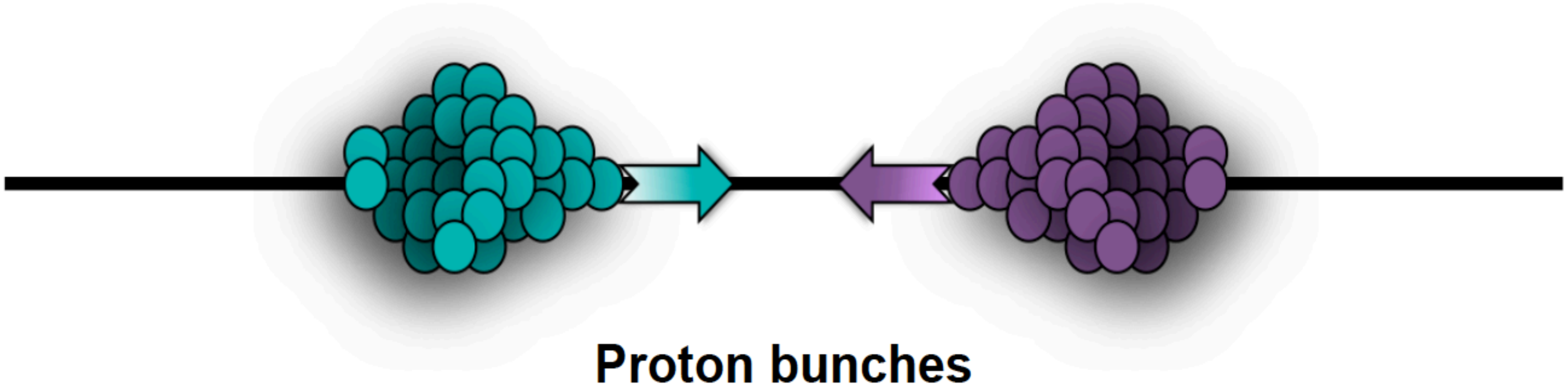


A crossing of the two LHC proton beams at an interaction point



Relative beam sizes around Interaction Point 1 (IP1) ATLAS in collision

WHAT IS AN EVENT?

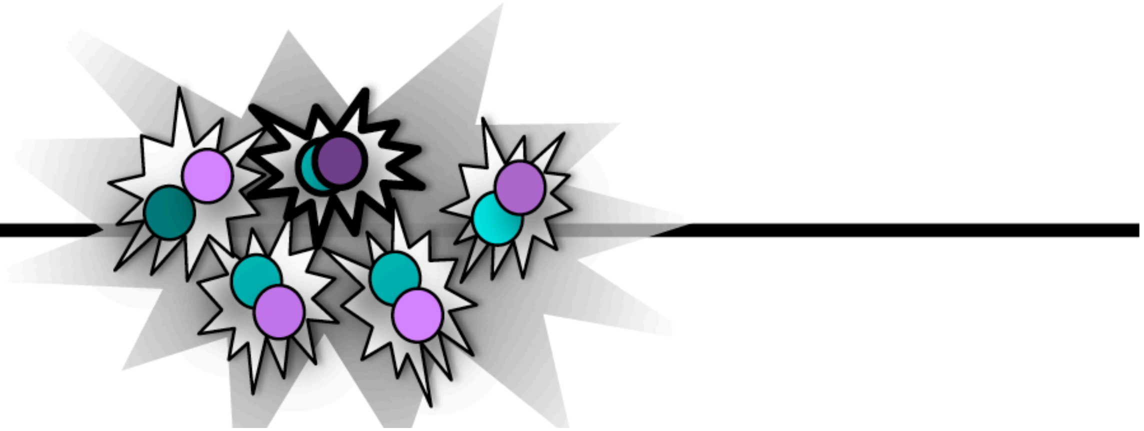


Proton bunches
>10^{11} protons/bunch

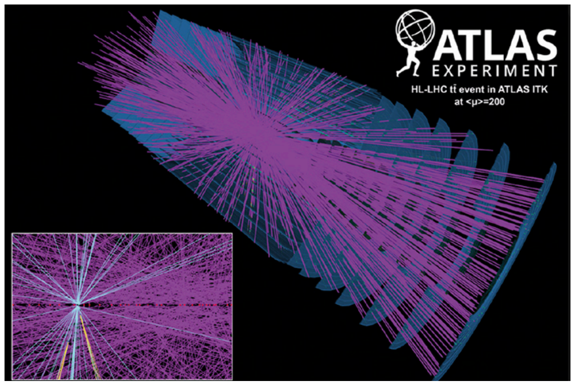
Colliding at 13/14 TeV and at 40 MHz in Run2 (2015-2018)
Collided at 7/8 TeV and at 20 MHz in Run1 (2010-2013)
commissioned at 2.76 TeV before Run1

PILE-UP

Proton bunches
>10^{11} protons/bunch
(colliding at **~40MHz in run2)**



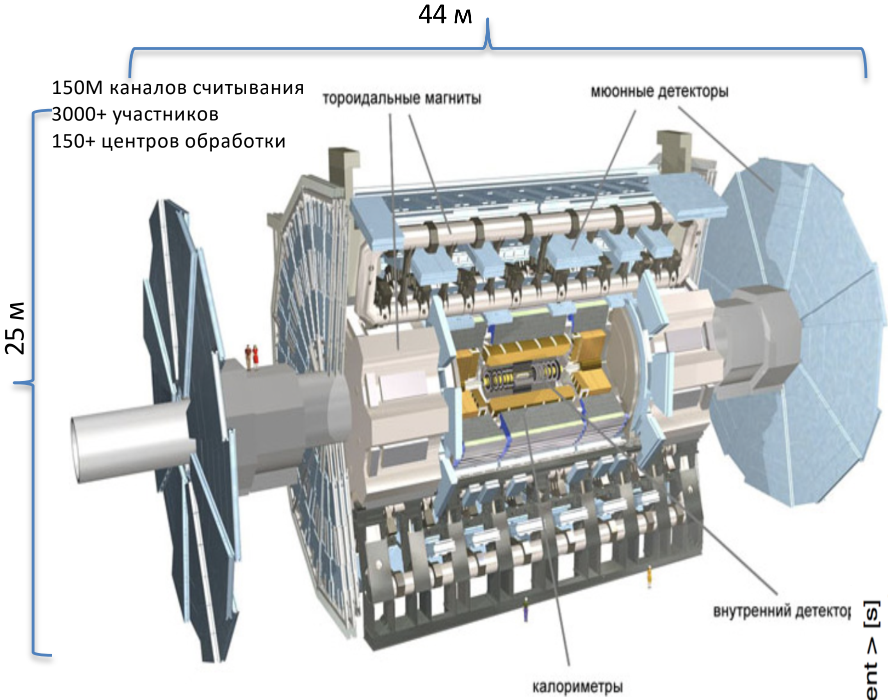
ATLAS: simulation for HL-LHC with 200 vertices



~20 p-p collisions / bunch crossing (Run I)
~40 p-p collisions / bunch crossing (Run II)

Increased complexity due to much **higher pile-up** and higher trigger rates will bring several challenges to reconstruction algorithms

Эксперимент ATLAS на Большом Адронном Коллайдере



ATLAS и CMS имеют следующие основные направления исследований :

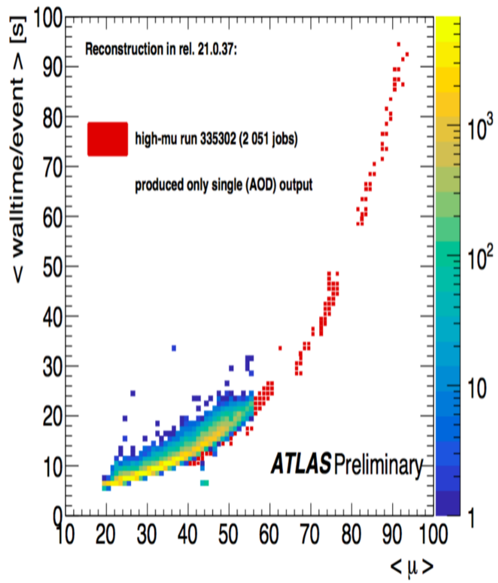
Использование бозона Хиггса как инструмент для новых открытий;

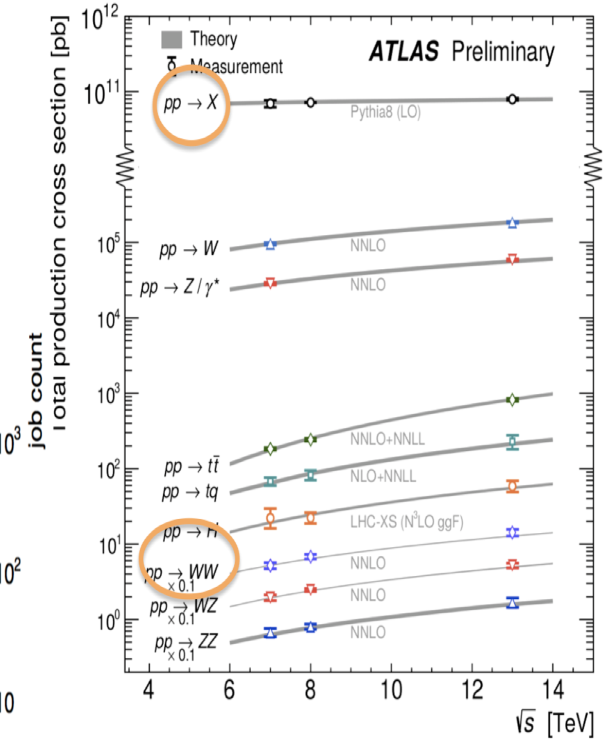
Поиск темной материи;

Поиск новой физики : частиц, взаимодействий, физических законов.

Объем данных ATLAS

Время реконструкции события





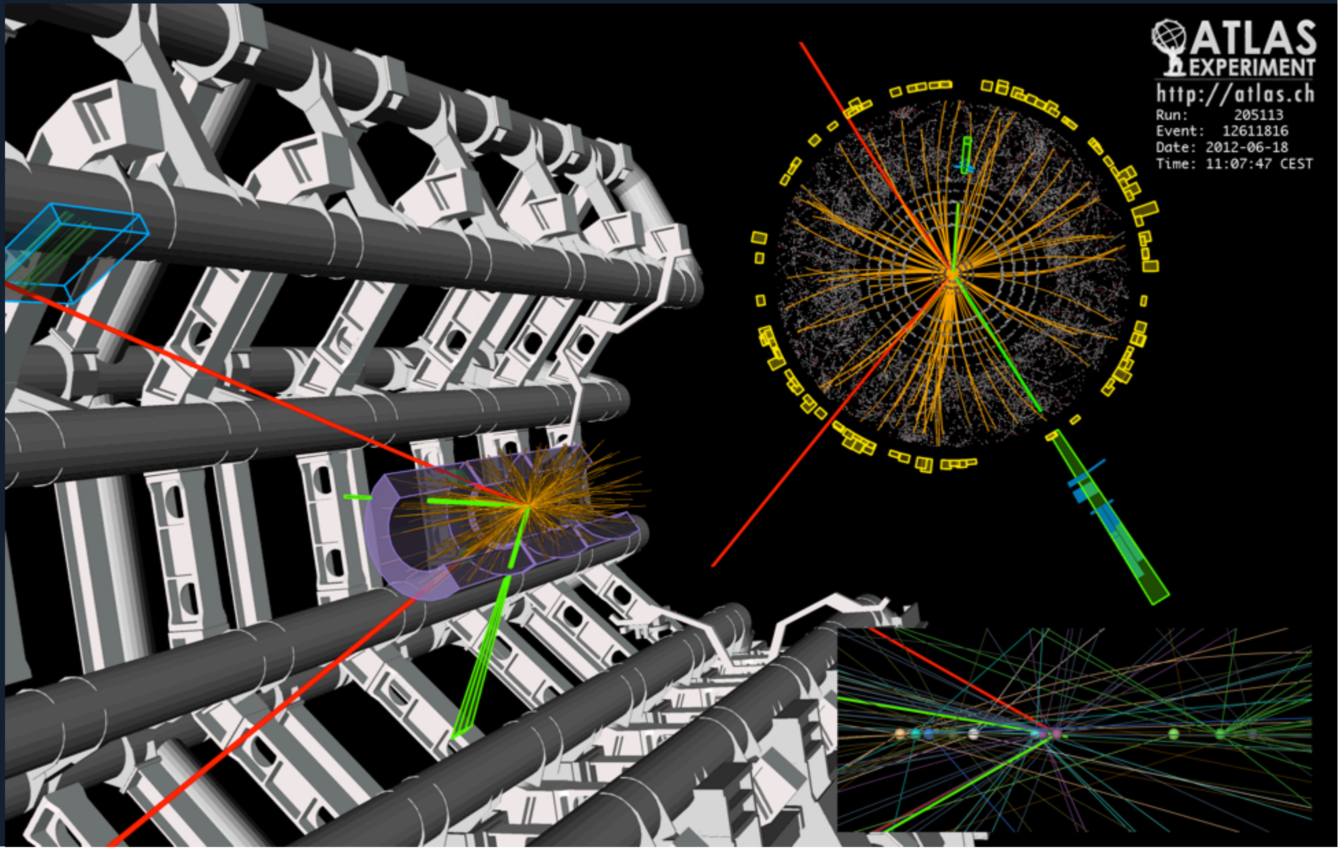
What is this data?

- Raw data:
  - Was a detector element hit?
  - How much energy?
  - What time?

- Reconstructed data:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information
  - …

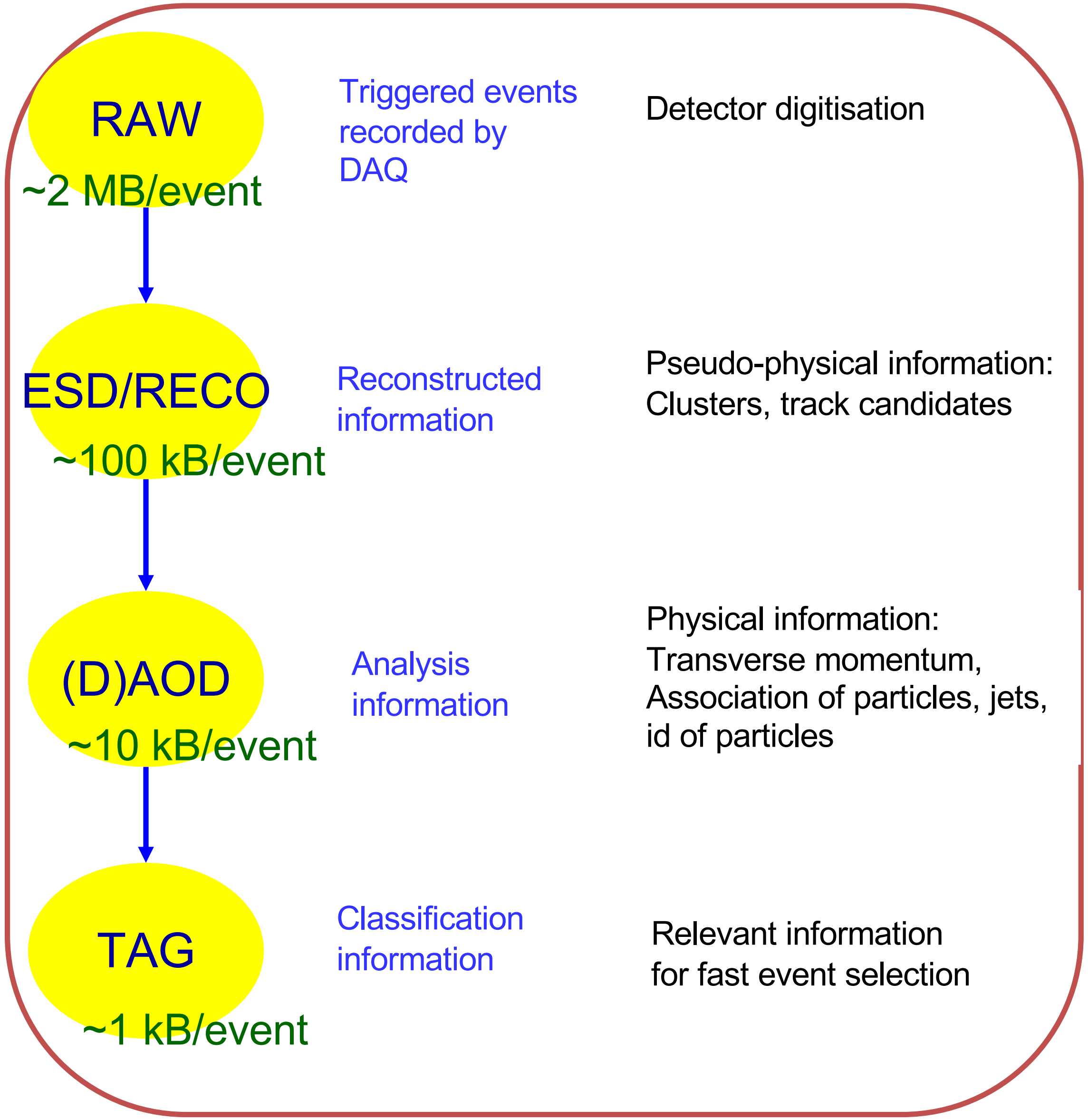
- 150 Million sensors deliver data … ~ 40 Million times per second

- Up to 6 GB/s to be stored and analysed after filtering

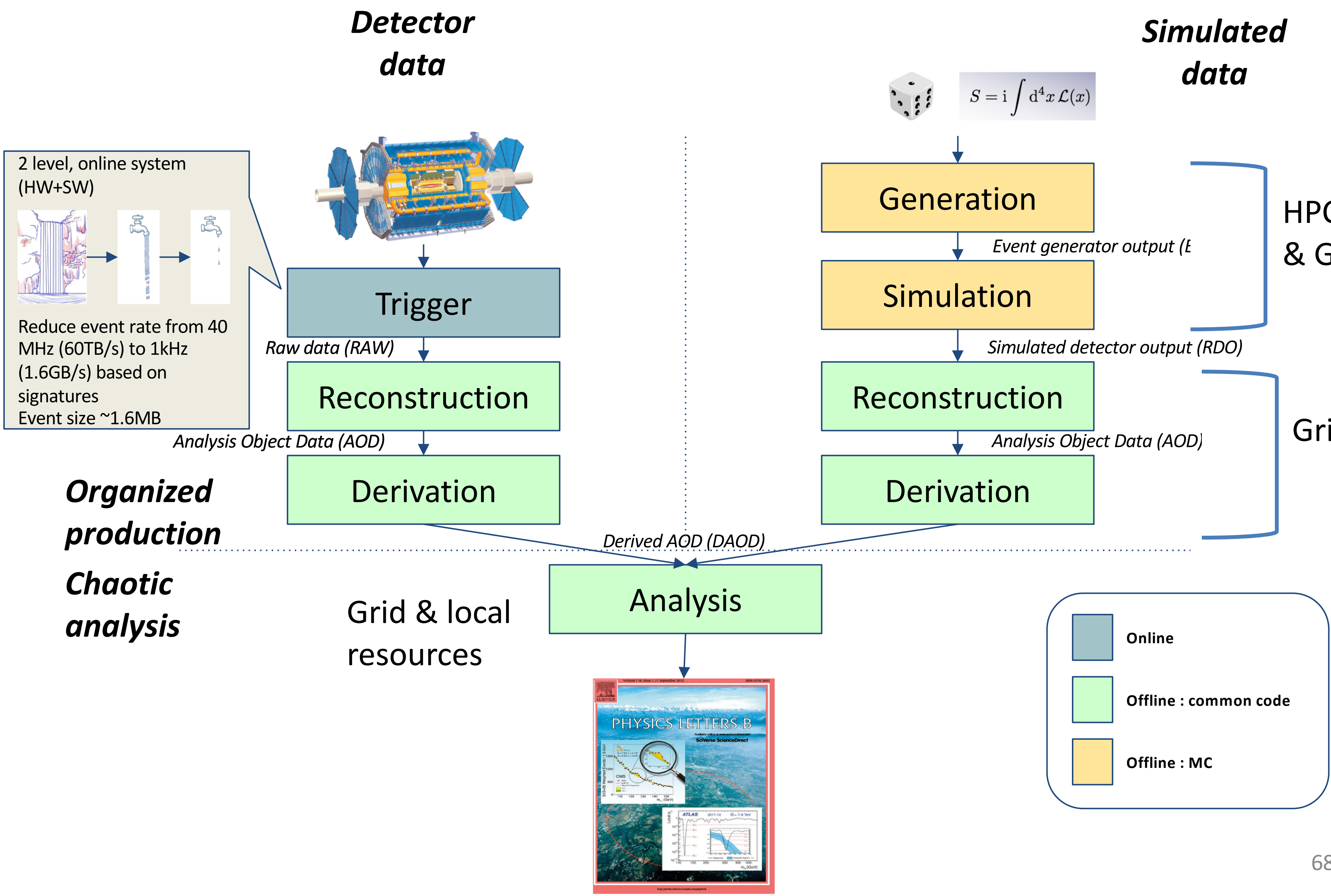


Data and Algorithms

- HEP data are organized as *Events* (particle collisions)
- Simulation, Reconstruction and Analysis programs process "one event at a time"
  - Events are fairly independent → Trivial parallel processing
- Event processing programs are composed of a number of Algorithms selecting and transforming "raw" event data into "processed" (reconstructed) event data and statistics
- *ATLAS reconstruction and simulation code 5M LOC*
  - *~70% C/C++, ~20% Python*
- *1000 software developers*



The data processing chain



The Data Processing Chain. Cont'd

Event Generation	Fundamental physics simulation of the proton collisions and resulting particle production; evolution into stable particles Processing rate O(0.01) seconds/event maximum	C++ Athena
Simulation	Simulation of the interaction of the particles with the detector external (Geant4) and internal (fast sim) O(10) minutes/event (Geant4), O(1) minutes/event (fast sim)	
Reconstruction	Re-building the collisions based on the detector (or simulated) data; conversion into the "event data model" for onward analysis O(30 seconds) / event	
Analysis	Processing of the reconstructed events by individual analysts; multi-variate analyses; final statistical analysis O(0.01) seconds/event	C++ Python ROOT ScyPy, etc

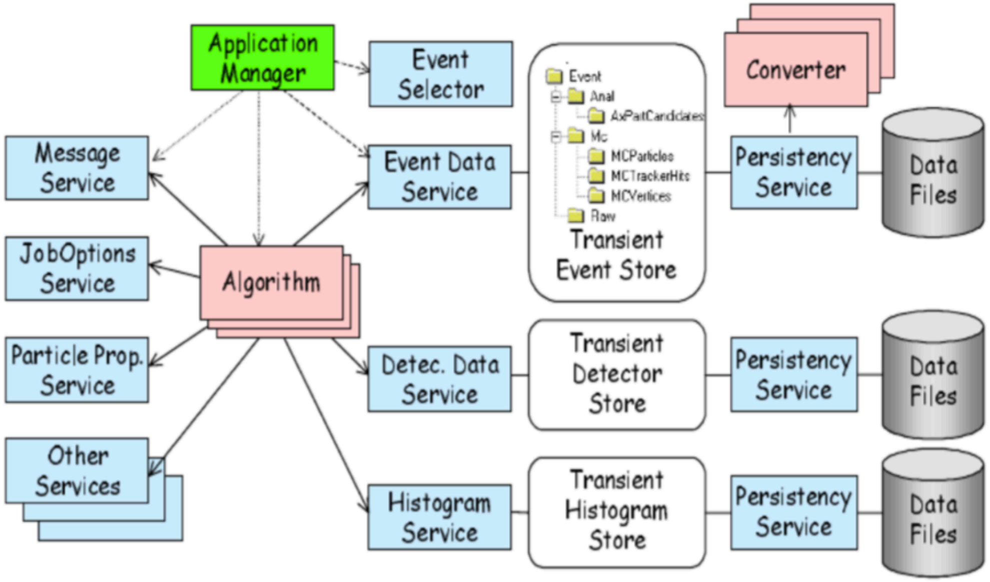
The Code

- The ATLAS offline code lives in 2 repositories
  - One reasonably small one that holds our basic build configuration code, and code for building external software needed by the offline code
  - One quite large one that holds all of the code itself
- ATLAS has a fair amount of code in our offline software
  - And this doesn't even count the amount of "common" HEP specific code that we absolutely need

```
Totals grouped by language (dominant language first):
cpp:        3892202 (69.77%)
python:     1164982 (20.88%)
xml:         342054 (6.13%)
sh:           61682 (1.11%)
fortran:      58333 (1.05%)
```

Software Framework : Athena / Gaudi

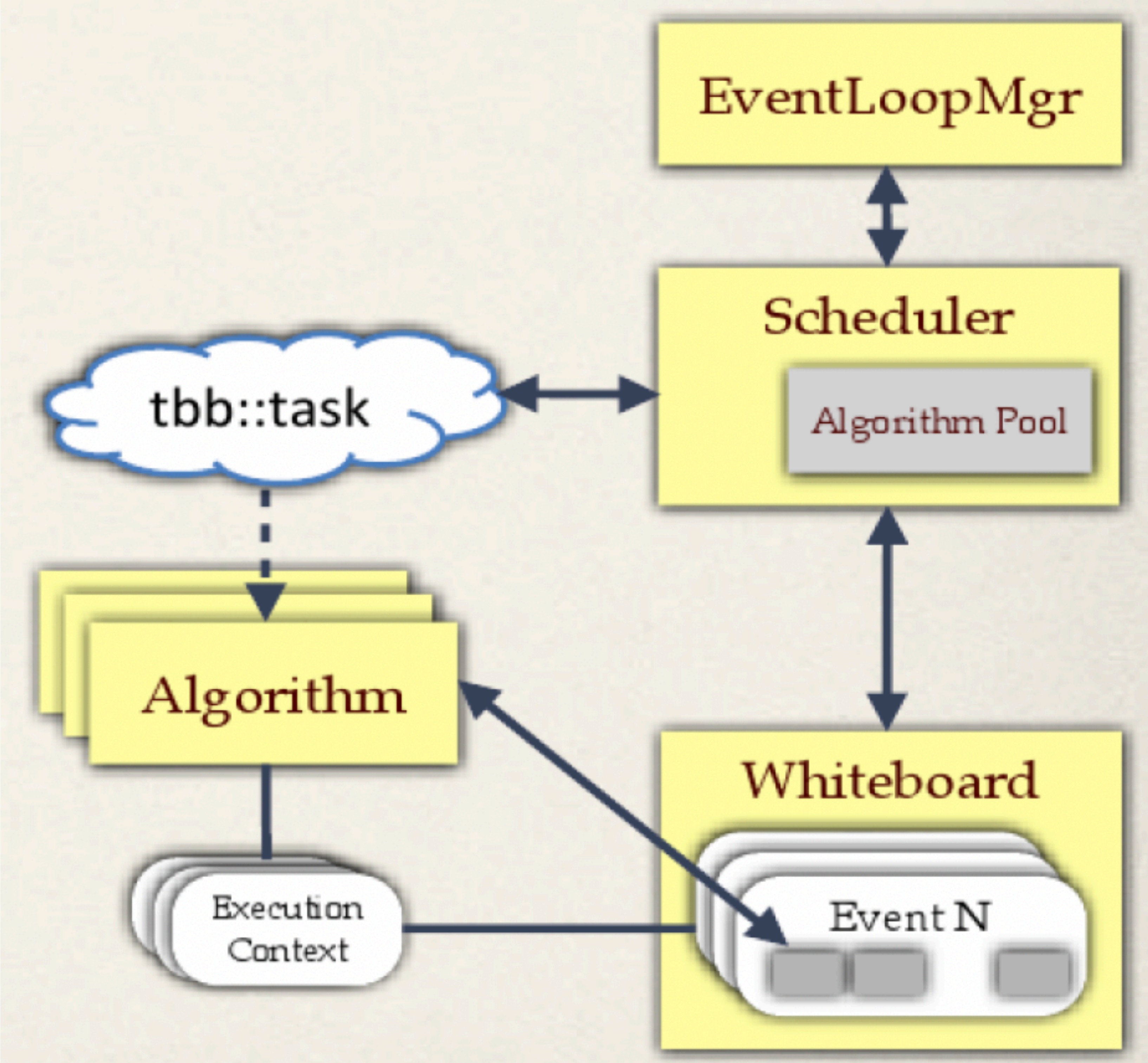
- Event processing
  - Algorithms run in predefined sequences
  - Algorithms produce and consume data via an Event Store
- Framework components
  - **Algorithm**
    - Main building block of the Event Loop called once per event
  - **AlgTool**
    - A plugin that helps an algorithm perform some action
  - **Service**
    - A plugin providing a common service to multiple components

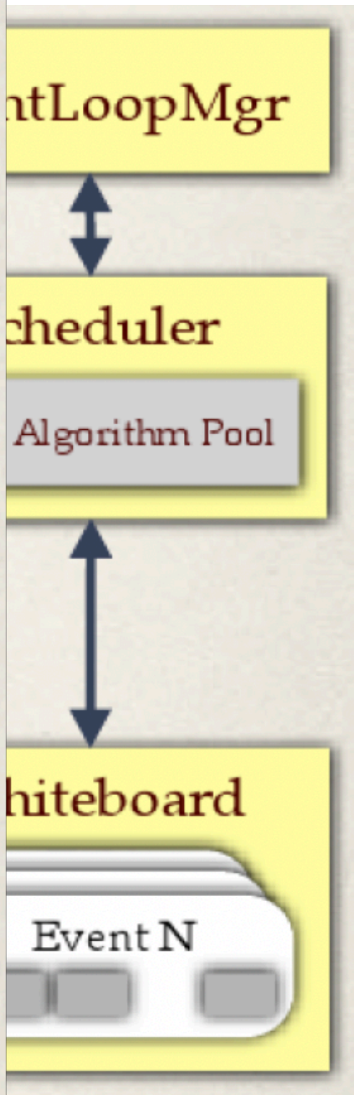


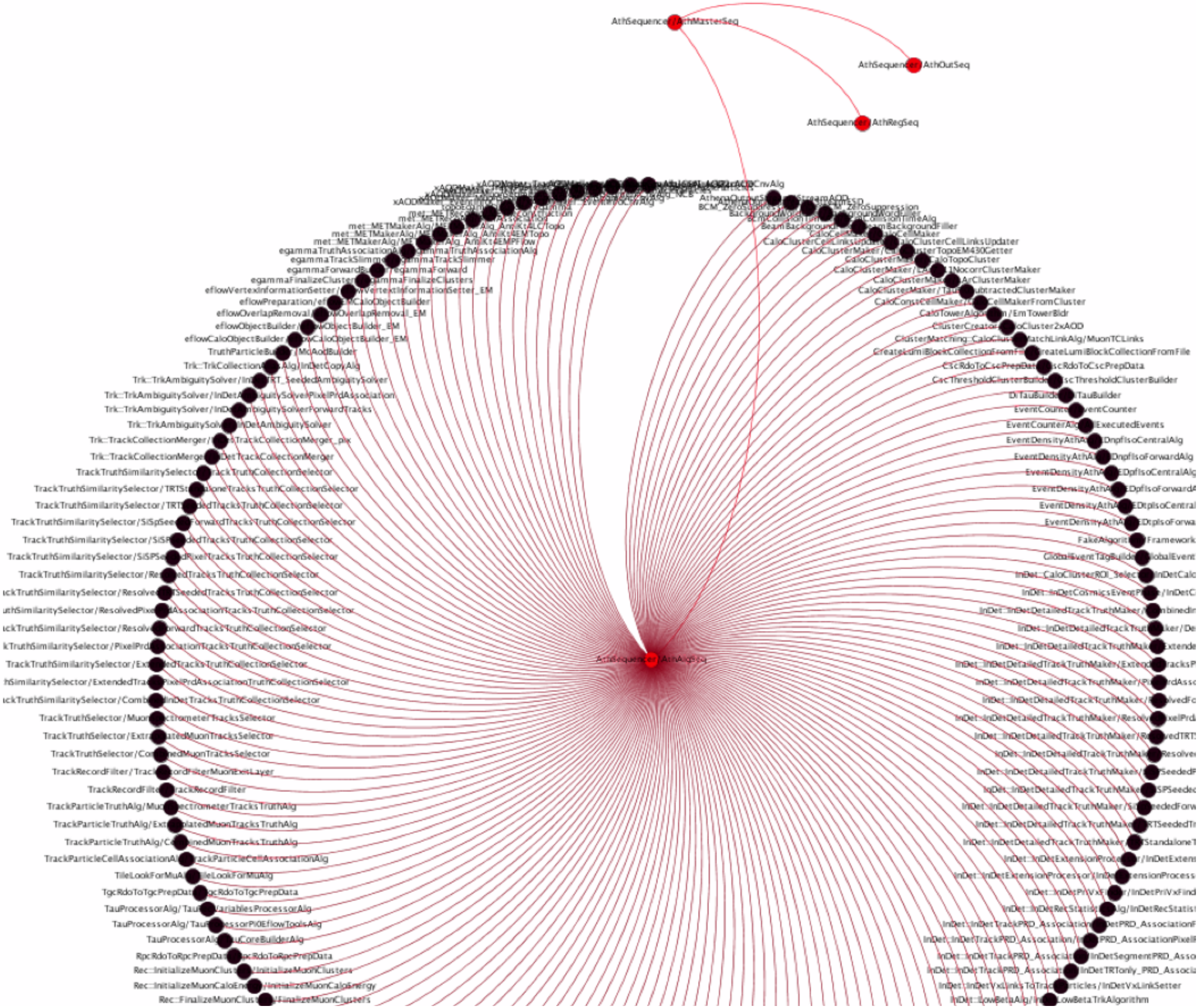
Athena/Gaudi framework was designed more than 15 years ago

Original execution model : a serial processing of events

work


EventLoopMgr
Scheduler
Algorithm Pool
tbb::task
Algorithm
Whiteboard
Event N
Execution
Context


Algorithm Pool
Event N



Experiment. RAW data...

fraction of RAW event

```
0x01e84c10:  0x01e8 0x8848 0x01e8 0x83d8 0x6c73 0x6f72 0x7400 0x0000
0x01e84c20:  0x0000 0x0019 0x0000 0x0000 0x01e8 0x4d08 0x01e8 0x5b7c
0x01e84c30:  0x01e8 0x87e8 0x01e8 0x8458 0x7061 0x636b 0x6167 0x6500
0x01e84c40:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c50:  0x01e8 0x8788 0x01e8 0x8498 0x7072 0x6f63 0x0000 0x0000
0x01e84c60:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c70:  0x01e8 0x8824 0x01e8 0x84d8 0x7265 0x6765 0x7870 0x0000
0x01e84c80:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84c90:  0x01e8 0x8838 0x01e8 0x8518 0x7265 0x6773 0x7562 0x0000
0x01e84ca0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cb0:  0x01e8 0x8818 0x01e8 0x8558 0x7265 0x6e61 0x6d65 0x0000
0x01e84cc0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cd0:  0x01e8 0x8798 0x01e8 0x8598 0x7265 0x7475 0x726e 0x0000
0x01e84ce0:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84cf0:  0x01e8 0x87ec 0x01e8 0x85d8 0x7363 0x616e 0x0000 0x0000
0x01e84d00:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d10:  0x01e8 0x87e8 0x01e8 0x8618 0x7365 0x7400 0x0000 0x0000
0x01e84d20:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d30:  0x01e8 0x87a8 0x01e8 0x8658 0x7370 0x6c69 0x7400 0x0000
0x01e84d40:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d50:  0x01e8 0x8854 0x01e8 0x8698 0x7374 0x7269 0x6e67 0x0000
0x01e84d60:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d70:  0x01e8 0x875c 0x01e8 0x86d8 0x7375 0x6273 0x7400 0x0000
0x01e84d80:  0x0000 0x0019 0x0000 0x0000 0x0000 0x0000 0x01e8 0x5b7c
0x01e84d90:  0x01e8 0x87c0 0x01e8 0x8718 0x7377 0x6974 0x6368 0x0000
```

- More than 300K such words in each event, corresponding to the full data from all the detector components.
- Data size: 1-1.5MB / event depending on the compression. Pretty consistent between ATLAS and CMS.
- **Challenge: make sense out of all these numbers!!**

"Address" :
- which detector element took the reading

"Value(s)" :
- what the electronics wrote out

Reconstructed Data.

Event 1

Nch (charged tracks) :
2

Pcha
(Momentum of each track):
{{"-7.65698","42.9725","14.3404"},
{" 7.54101","-42.1729","-14.0108"}}
px py pz

Qcha
(Charge of each track):
{-1,1}

Event 2

Nch (charged tracks) :
3

Pcha
(Momentum of each track):
{{"-12.9305","12.2713","40.5615"},
{" 12.2469","-11.606","-38.7182"},
{"0.143435","-0.143435","-0.497444"}}
px py pz

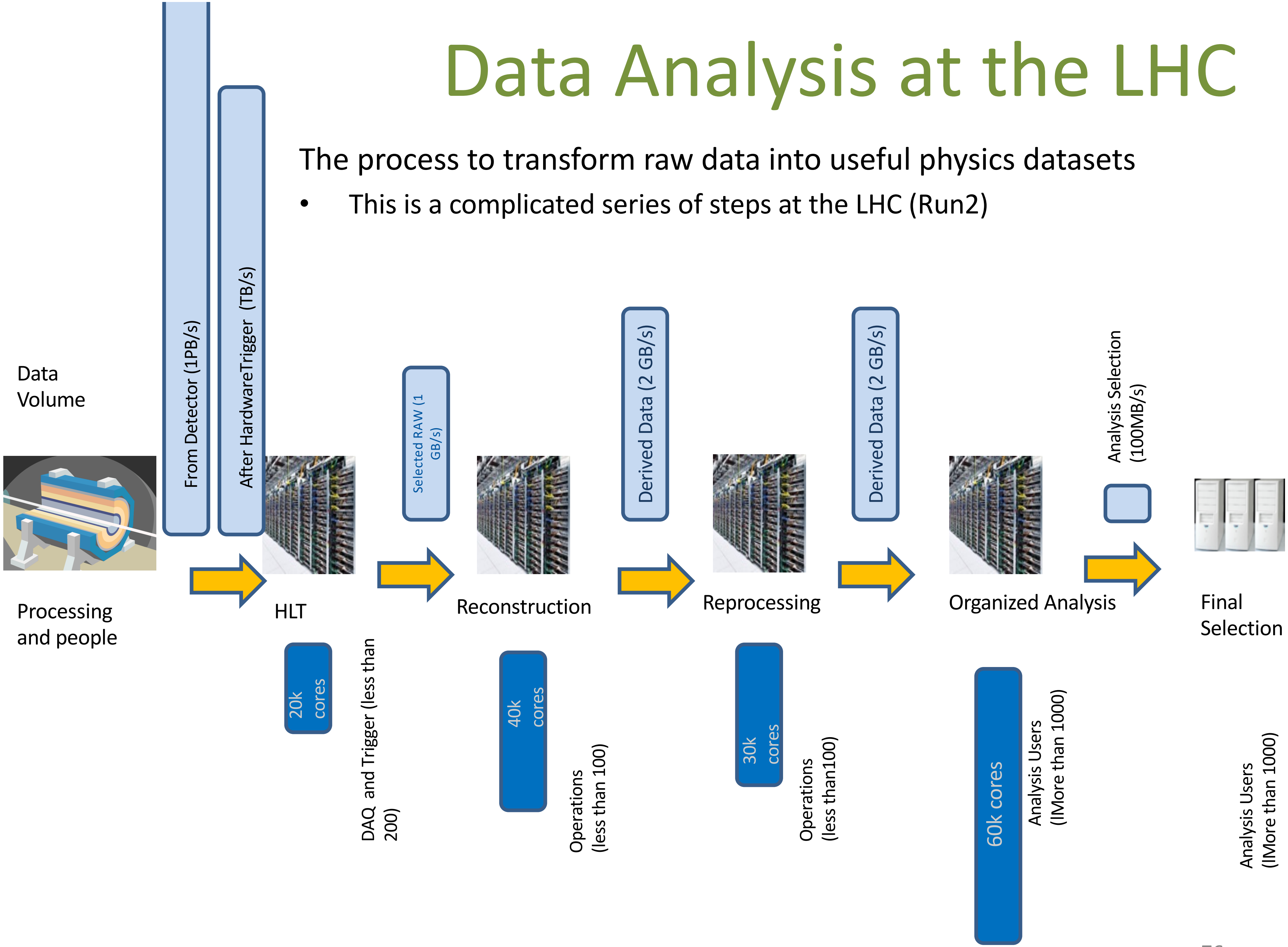
Qcha
(Charge of each track):
{-1,1,-1}

File A

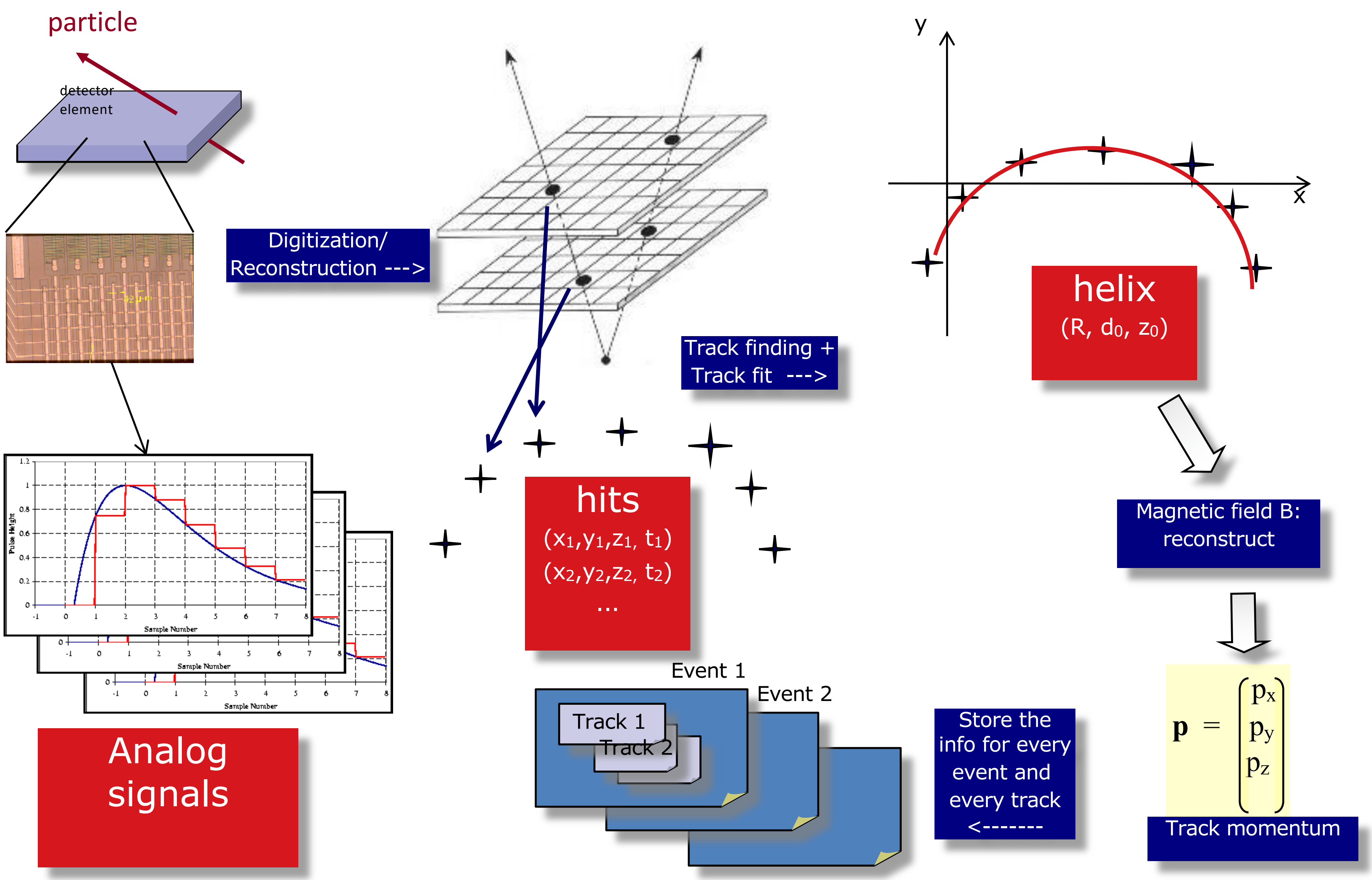
Data Analysis at the LHC

The process to transform raw data into useful physics datasets

- This is a complicated series of steps at the LHC (Run2)

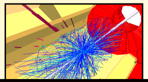


Data reduction/abstraction



Making the connection

Reality



Raw Data — The imperfect measurement of a (set of) interactions in the detector

↓

Reconstructed Events — A unique happening:
eg. Run 23458, event 1345
which contains a Z → $\mu^+\mu^-$ decay

↓

Analysis : We "confront theory with experiment" by comparing the measured quantity (observable) with the prediction.

Observables — cross sections (probabilities for interactions), branching ratios (BR), ratios of BRs, specific lifetimes, ...

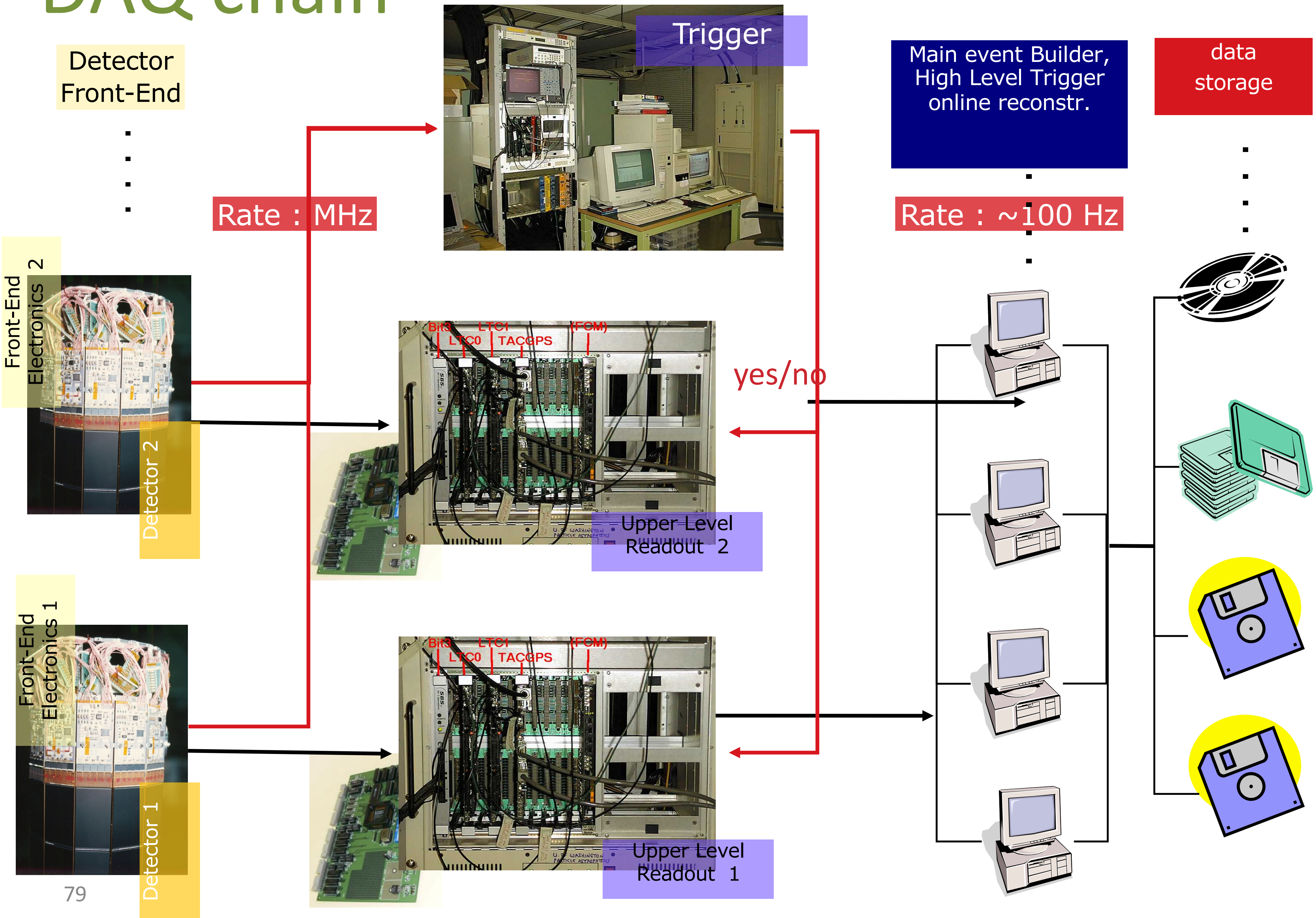
↑

Theory

$$S = \mathrm{i}\int \mathrm{d}^4x\,\mathcal{L}(x)$$

A small number of general equations, with some parameters (poorly or not known at all)

DAQ chain


Detector Front-End
Trigger
Main event Builder, High Level Trigger online reconstr.
data storage
Rate : MHz
Rate : ~100 Hz
Front-End Electronics 2
Detector 2
Front-End Electronics 1
Detector 1
yes/no
Upper Level Readout 2
Upper Level Readout 1

TRIGGER MENUS

Trigger	Typical offline selection	L1 Peak Rate (kHz) L_{peak}=7e33/cm^{2}s	EF Avg. Rate (Hz) $L_{avg.}$=5e33/cm^{2}s
Single leptons	Single iso μ, p_T >25 GeV	8	45
	Single iso e, p_T >25 GeV	17	70
Two leptons	Two μ's, each p_T >15 GeV	1	5
	Two μ's, p_T >20, 10 GeV	8	8
	Two e's, each p_T >15 GeV	6	8
	Two e's, p_T >25, 10 GeV	17	5
	Two τ's, p_T >45, 30 GeV	12	12
Two photons	Two γ's, each p_T >25 GeV	6	10
	Two γ's, p_T >40, 30 GeV	6	7
Single jet	Jet (R = 0.4), p_T >360 GeV	2	5
	Jet (R = 1.0), p_T >470 GeV		2
E_T^{miss}	E_T^{miss} >150 GeV	2	17
Multi-jets	4 jets, each p_T >85 GeV	1	8
	5 jets, each p_T >60 GeV		2
	6 jets, each p_T >50 GeV		4
b−jets	4 jets, each p_T >50 GeV out of which one is b−tagged	1	4
Total		< 75	400

STREAMING

- Streaming is based on trigger decisions at all stages
- The Raw Data physics streams are generated at the HLT output level

Debug Streams
events for which a trigger decision has not been made, because of failures in parts of the online system

Physics Streams
data for physics analyses

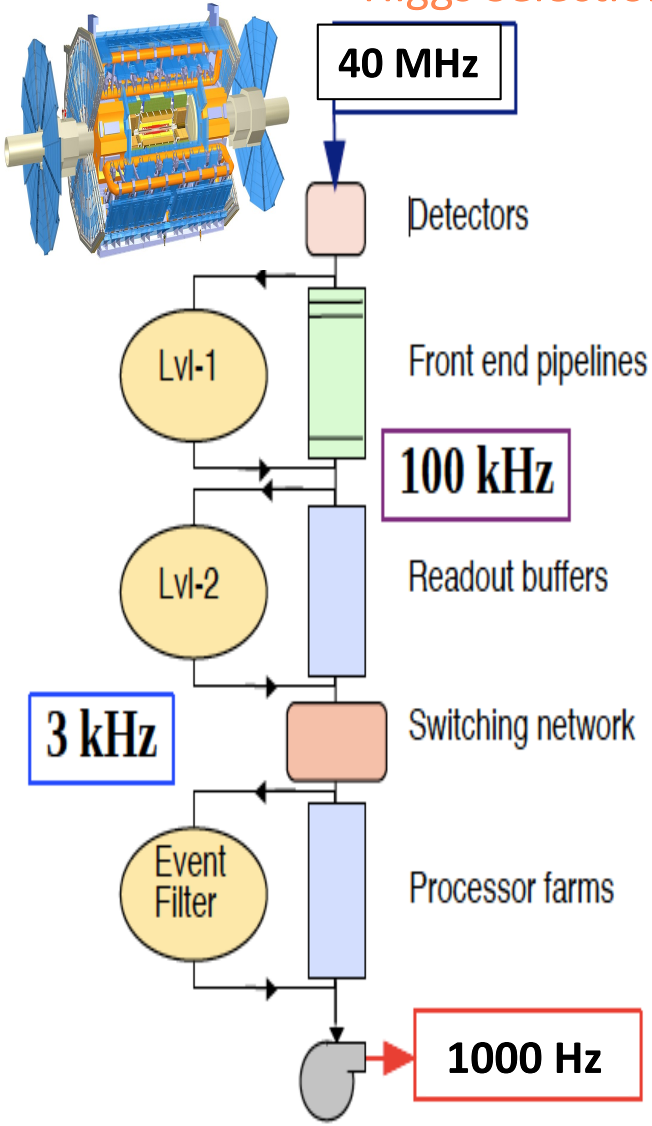
Express Stream
full events for fast reconstruction

Calibration Streams
events delivering the minimum amount of information for detector calibrations at high rate

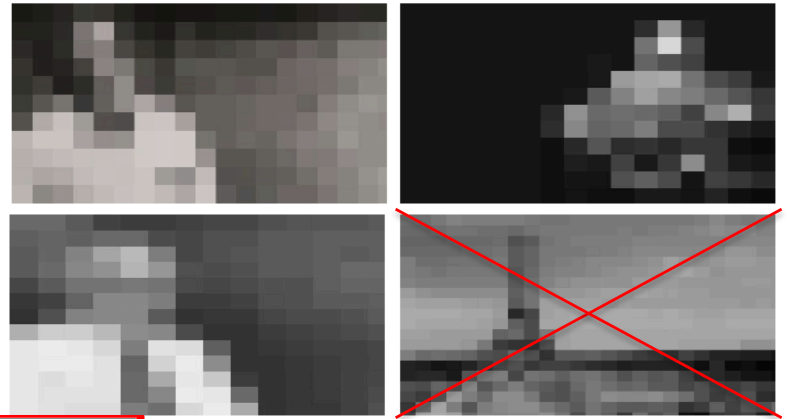
HLT – High Level Trigger Farm @CERN

HEP Online. Reduce the data volume in stages.

Higgs Selection using the Trigger

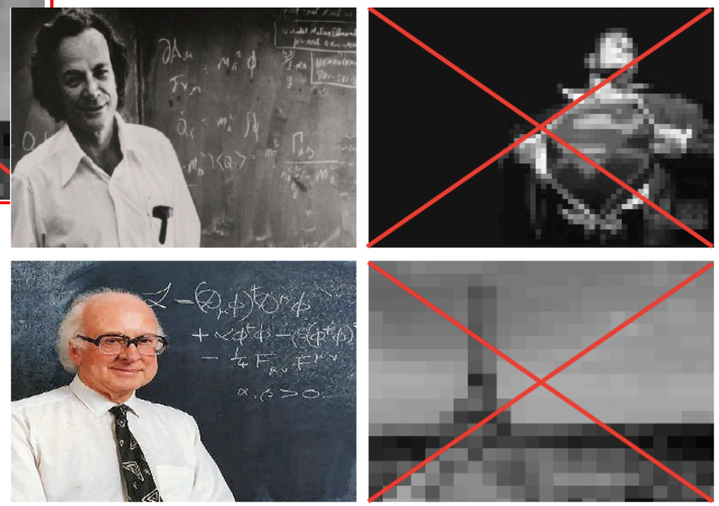


Level 1:
Not all information available, coarse granularity



Level 2:
Reconstruct events
Improved ability to reject events

Level 3:
High quality reconstruction algorithms, using information from all detectors



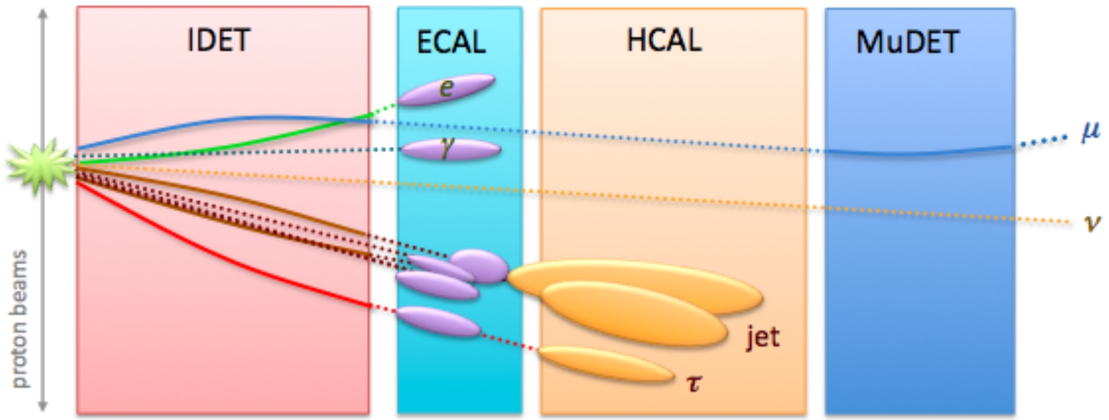
1000 Hz. Run2 ATLAS RAW data rate to tape

Классификация потоков заданий современного эксперимента в области Физики Высоких Энергий и Ядерной Физики.

- Моделирование методом Монте-Карло
- Обработка и переобработка данных эксперимента. Обработка данных для системы отбора событий “высшего” уровня (триггер HLT).
- Обработка, фильтрация и анализ данных, проводимые физическими группами.
- Физический анализ данных.
- Обработка данных “поездом” и “постоянная обработка” данных.

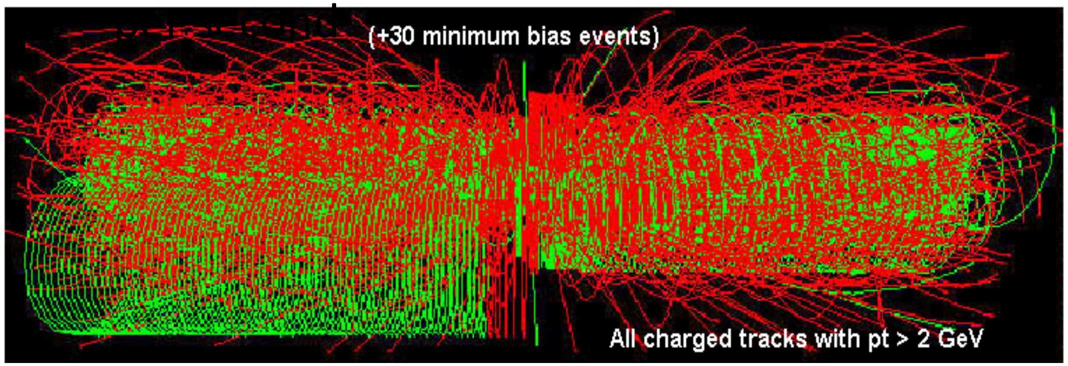
Reconstruction

- Detector reconstruction
  - Tracking
    - finding path of charged particles through the detector
  - Calorimeter reconstruction
    - finding energy deposits in calorimeters from charged and neutral particles
- Combined reconstruction
  - Electron/Photon identification
  - Muon identification
  - Jet finding
- Calibrations and alignments applied at nearly every step



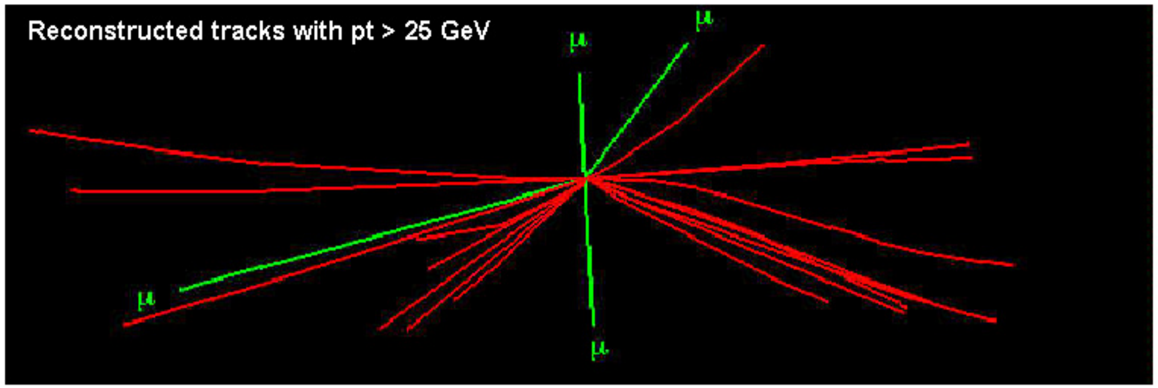
Reconstruction Goals.

- 800,000,000 proton-proton interactions



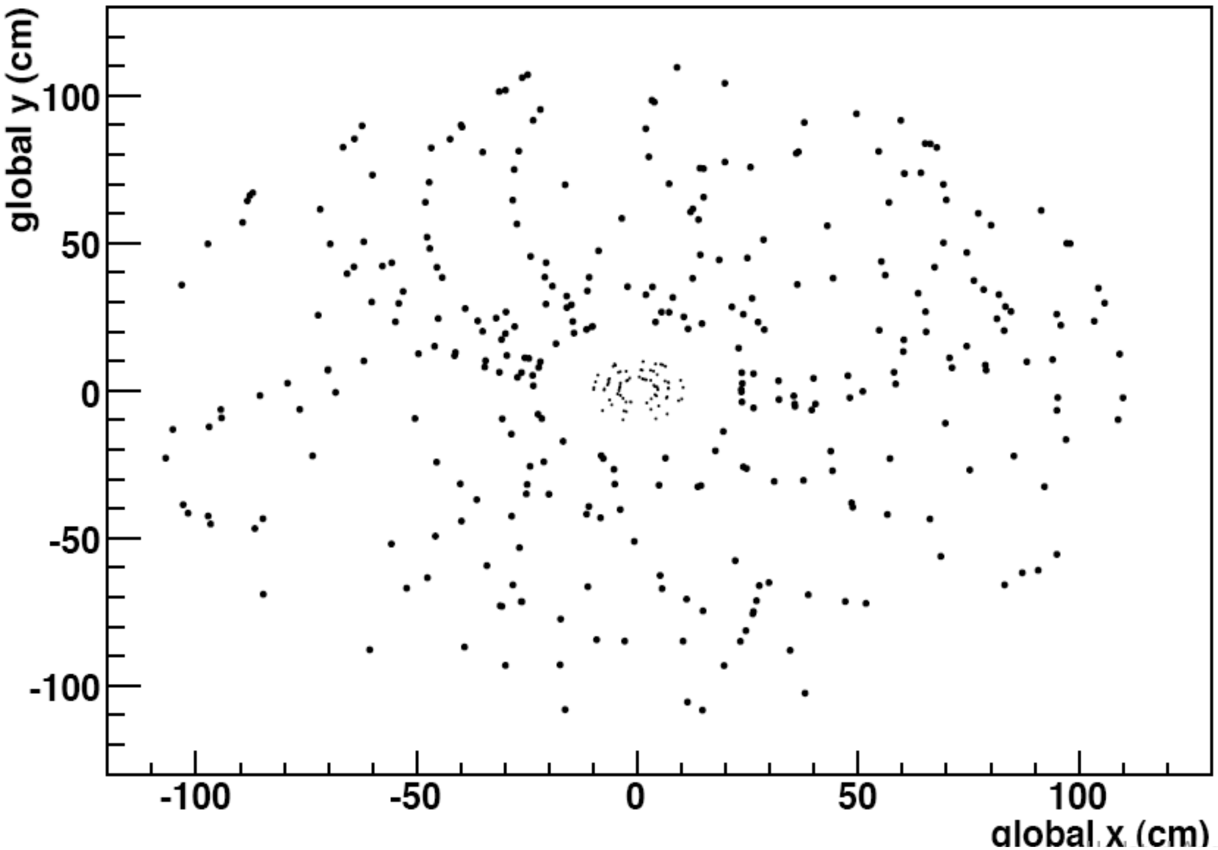
- High efficiency
- Good resolution
- Low fake rate
- Robust against detector problems
  - Noise
  - Dead regions of the detector
- Be able to run within the computing resource limitations
  - CPU time per event
  - Memory use

We are looking for this "signature"



Reconstruction. Track Finding.

- Track Finding
  - Track finding very important for analysis
  - Tracks are used directly in the reconstruction of
    - Electrons
    - Muons
      - And to a lesser extent in Tau, Jet and photon reconstruction
  - For reconstructed tracks we know
    - Momentum
      - straighter the track the higher momentum it is
    - Charge
    - Point of closest approach to the interaction point



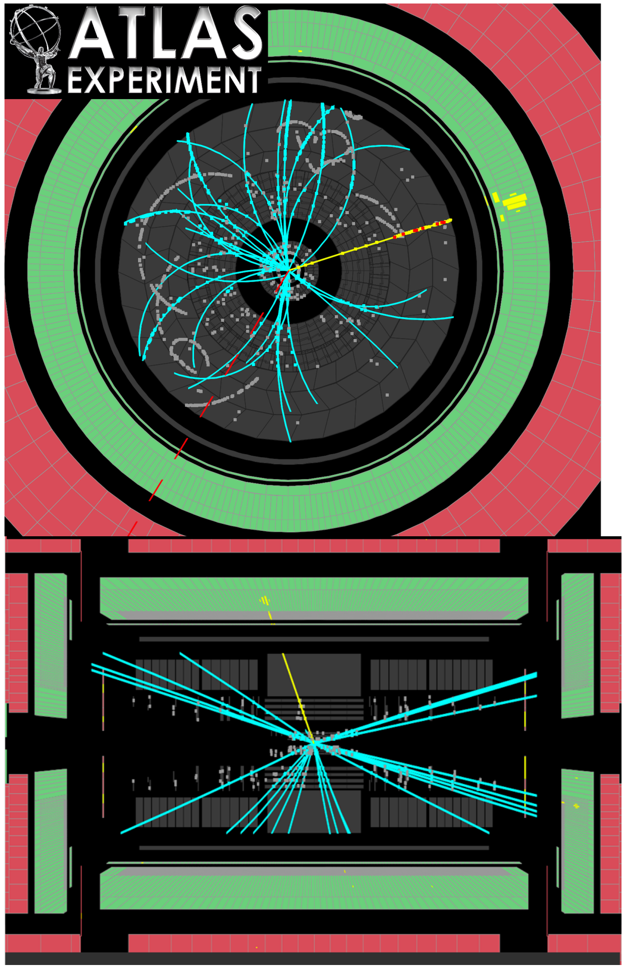
Tracking by eye – Can you find the 50 GeV Track?

Physics Objects Reconstruction.

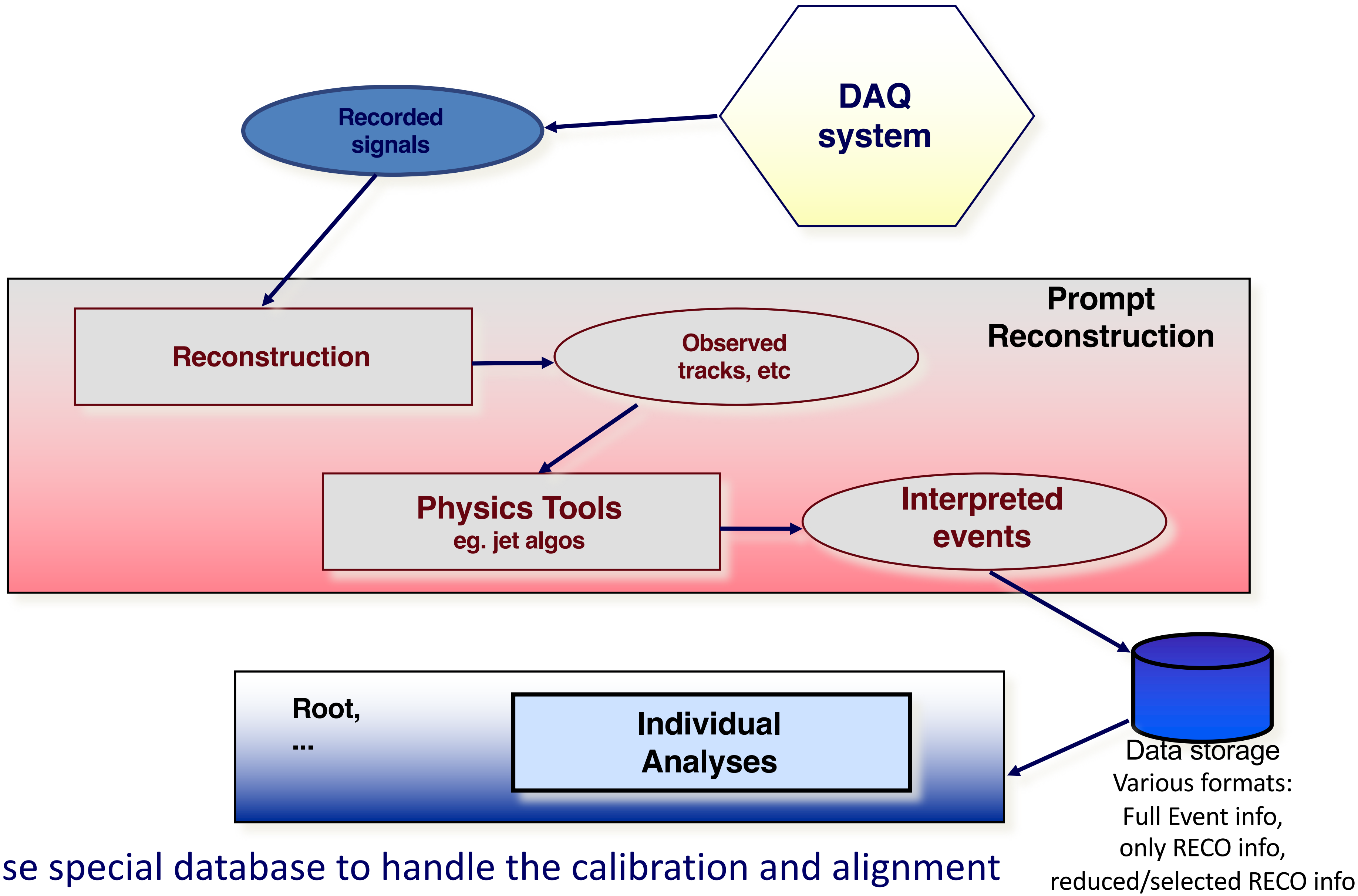
- Electron/Photon Identification
  - Electron/Photon reconstruction takes as input the tracks and calorimeter clusters already produced
  - Electron/Photon leave narrow clusters in the electromagnetic calorimeter
    - Apply selection on the cluster shape to reduce background from jets
  - Electron has track pointing at cluster
    - Requires aligning the calorimeter with the tracker
  - Photon has no track pointing at it
  - Final Electron momentum measurement can come from tracking or calorimeter information (or a combination of both)
    - Often have a final calibration to give the best electron energy
  - Often want isolated electrons
    - Require little calorimeter energy or tracks in the region around the electron

An electron in ATLAS

1. Narrow cluster in electromagnetic calorimeter
2. No energy in the hadronic calorimeter behind
2. High momentum track pointing at cluster

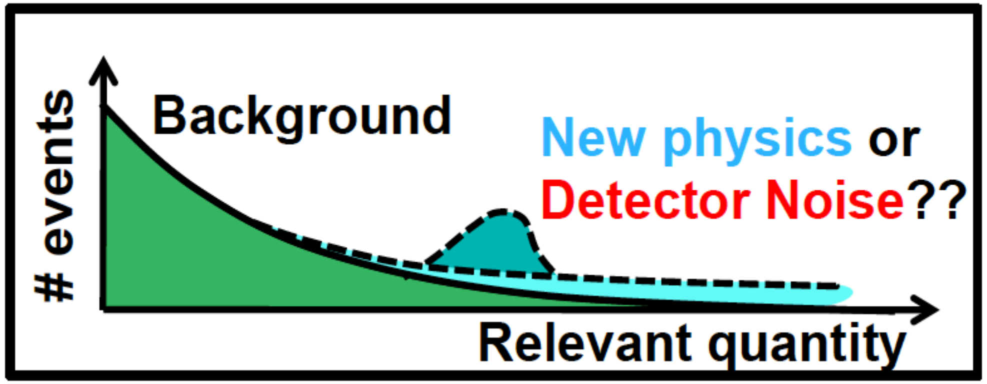


RECO flow



- Use special database to handle the calibration and alignment data needed in reconstruction

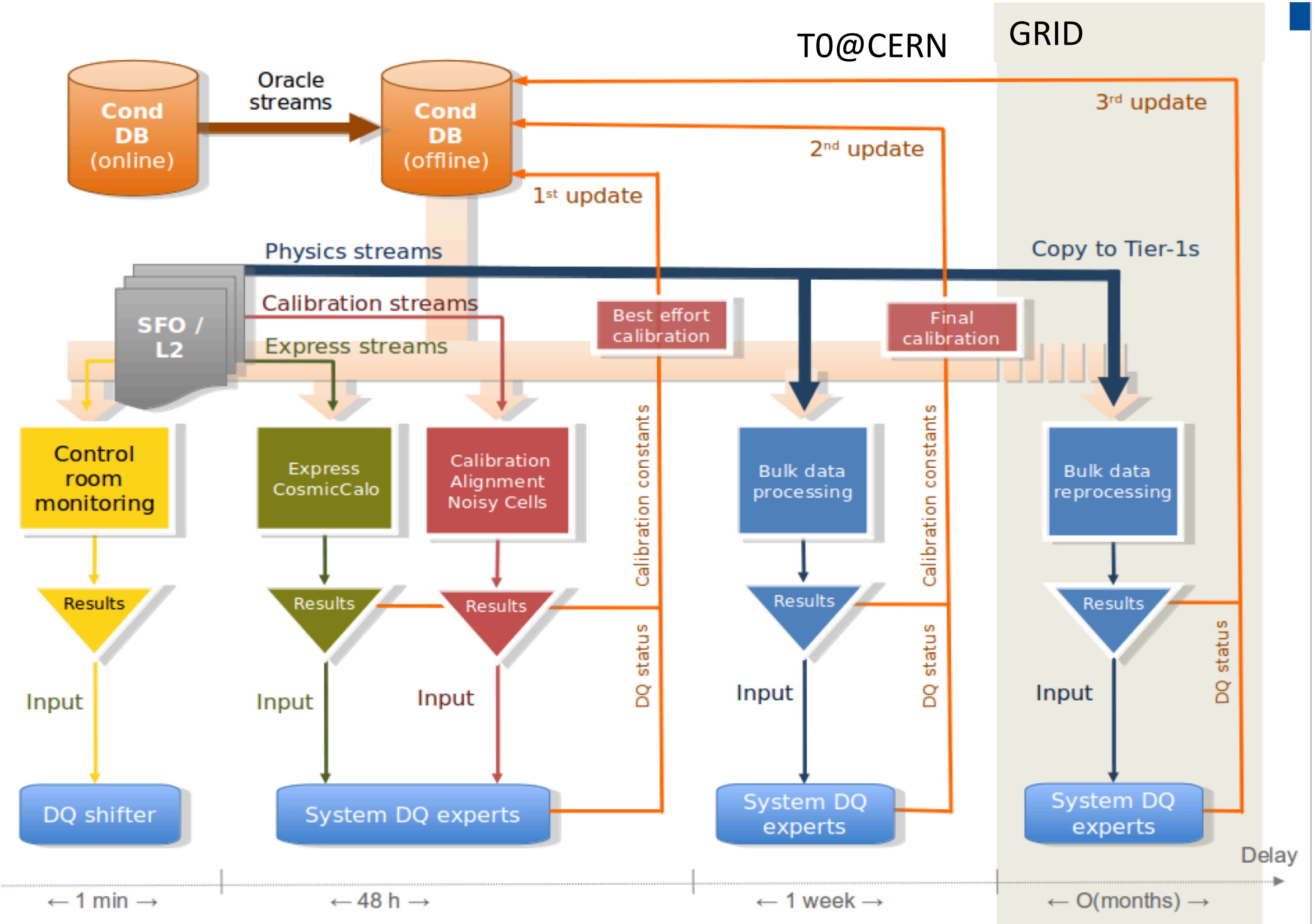
DATA QUALITY



The data we analyze has to follow norms of quality such that our results are trustable.

- **Online: Fast monitoring of detector performance during data taking, using dedicated stream, "express stream".**
- **Offline: More thorough monitoring at two instances:**
  - Express reconstruction; fast turn-around.
  - Prompt reconstruction: larger statistics.
- **What is monitored?**
  - Noise in the detector.
  - Reconstruction (tracks, clusters, combined objects, resolution and efficiency).
  - Input rate of physics.
  - All compared to reference histograms of data that has been validated as "good".

Data Quality



Simulation

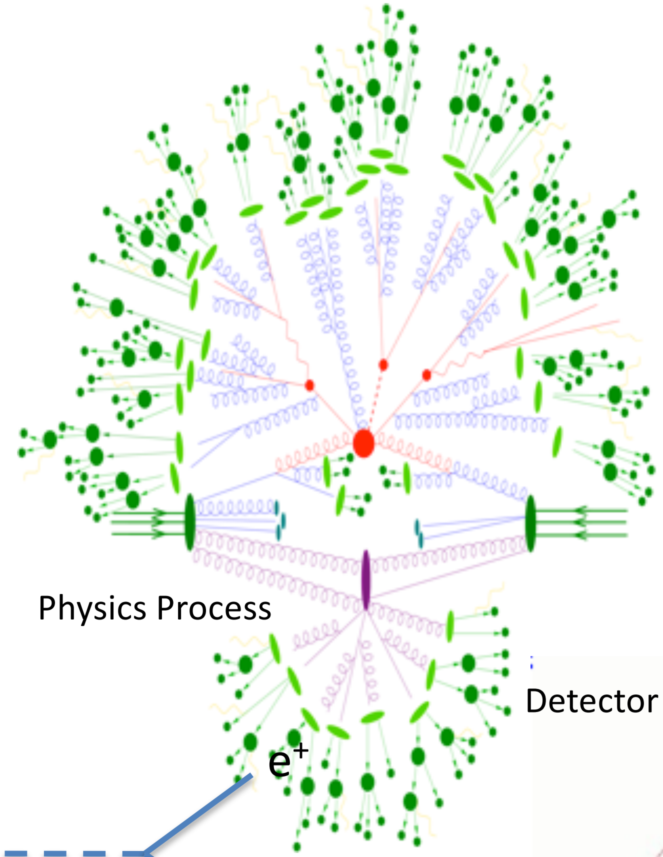
MONTE CARLO SIMULATION – WHY

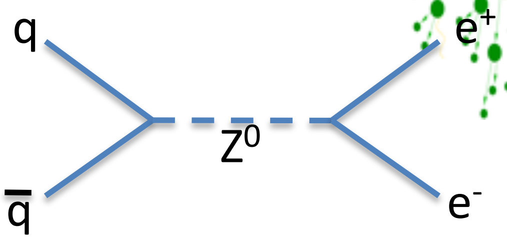
- **We only build one detector.**
  - How do we compromise physics due to detector design?
  - How would a different detector design affect measurements?
  - How does the detector behave to radiation?
- **In the detectors we only measure voltages, currents, times.**
  - It's an *interpretation* to say that such-and-such particle caused such-and-such signature in the detector.
  - Simulating the detector behavior we correct for inefficiencies, inaccuracies, unknowns.
- **We need a theory to tell us what we expect and to compare our data against.**
- **A good simulation is the way to demonstrate to the world that we understand the detectors and the physics we are studying.**

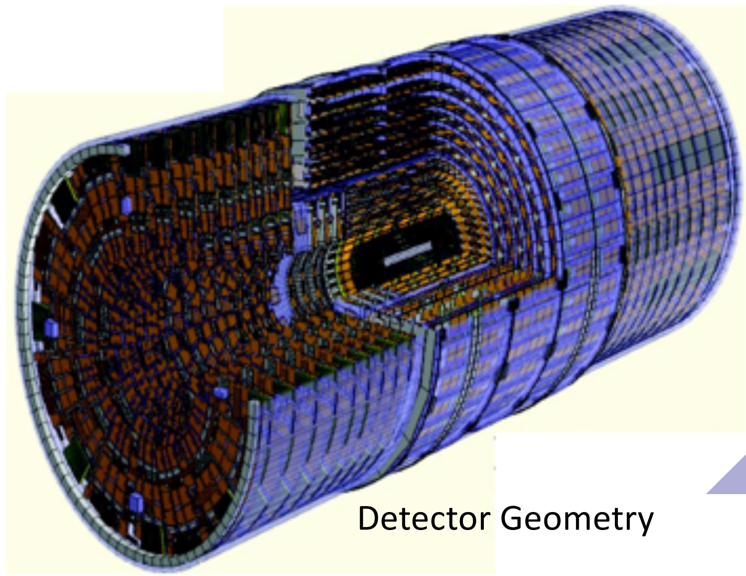
Simulation

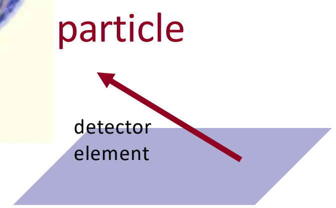
- Simulated data samples needed for
  - Designing experiments
  - Tuning analysis selections
  - Background estimation
  - Efficiency, resolution and fake-rate estimation
- To get best physics outputs from the experiment it is essential to have an accurate simulation of the detector
  - Lots of work goes into tuning the simulation to give best description of the data
  - Test beam studies from construction period of the detector used to tune simulation (test beam allows to study detector response to known particle types and momenta – e.g. 20GeV electrons)
- Very detailed simulation of the detector
  - Detailed description of the detector geometry
  - Accurate simulation of the detector electronics response
  - Include detector 'noise' in the simulation
- Keep the 'truth' information
  - Allows efficiency, resolution and fake rates to be estimated

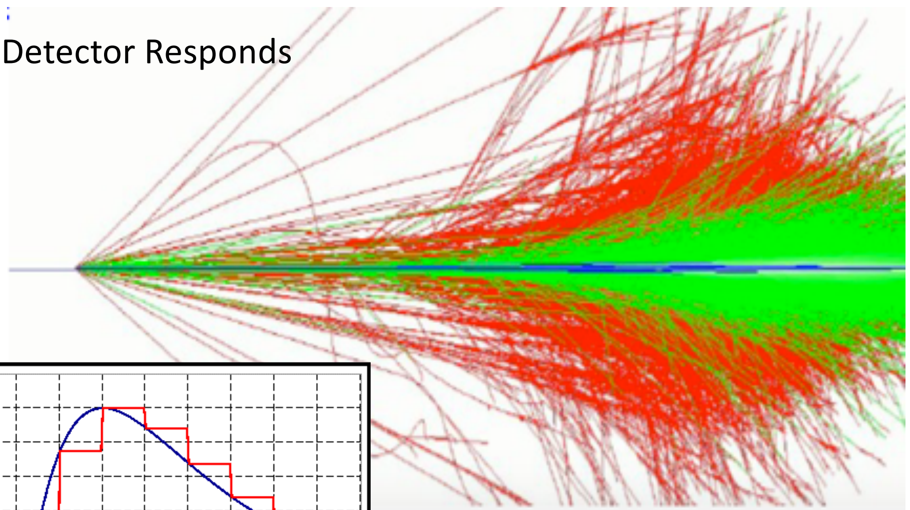
Simulation is complex at all levels

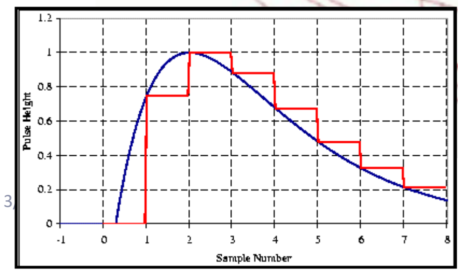








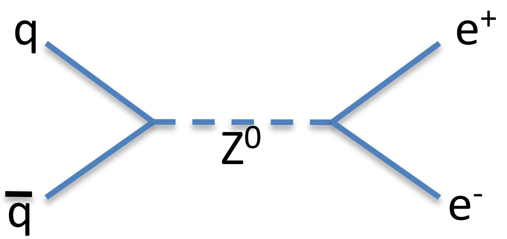




Simulation workflow

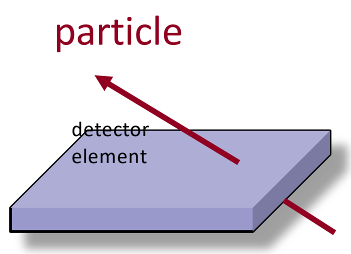
Physics simulation

Simulate the physics interaction (set in the simulation configuration)
Output of this part is the 4-vector's of the produced particles.
In this case the 4-vector's of the 2 electrons from the Z decay.



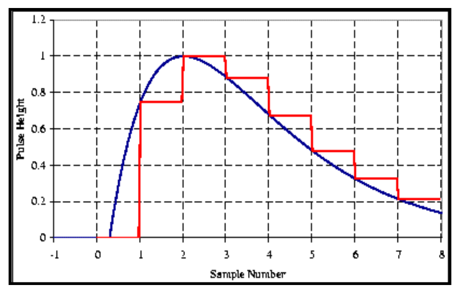
Detector Simulation

Simulate the propagation of the electrons through the detector.
Including:
-bending in the magnetic field
-leaving hits in the tracking detector elements
-interacting with the material in the detector
-interacting in the calorimeter (detailed description of the EM shower)



Electronics Simulation

Simulate the response of the detector elements to the 'hits' from the electron.
Simulate the voltage pulse on the detector and how the detector electronics works.
The output of this stage is very similar to the raw data from the detector.
(but we keep the truth information).



Detector simulation step is very CPU intensive. Requires huge computing resources.

MONTE CARLO PRODUCTION CHAIN

Event Generation
simulate the physics process.

Detector Simulation
simulate the interaction of the particles with the detector material.

Digitization
Translate interactions with detector into realistic signals.

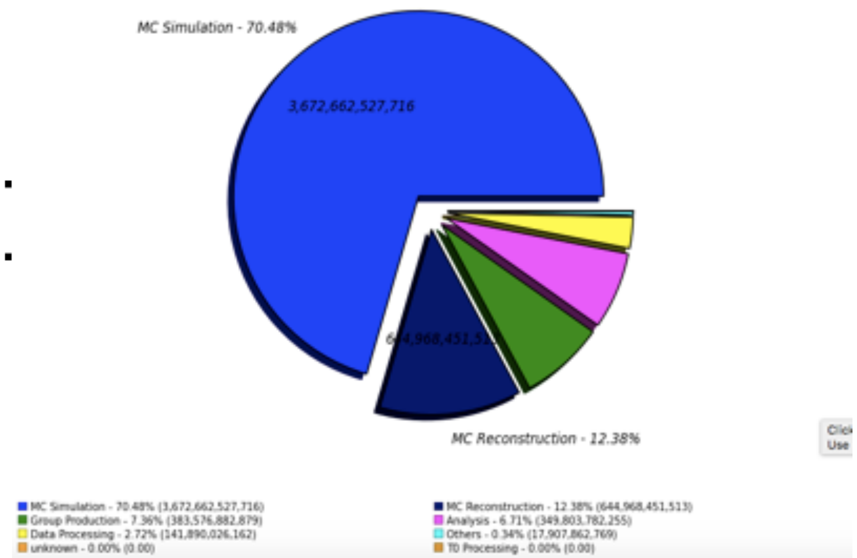
Reconstruction
Go from signals back to particles, as for real data.

How much processing time needed for each step?

From < 1s to a few hours / event.

From 1 to 10min / event

MC simulation 70.5% of ATLAS Computing Resources



All together, 70% of ATLAS computing resources are utilized to produce simulated events samples

MONTE CARLO PRODUCTION CHAIN

How much processing time needed for each step?

Event Generation
simulate the physics process.

. . .

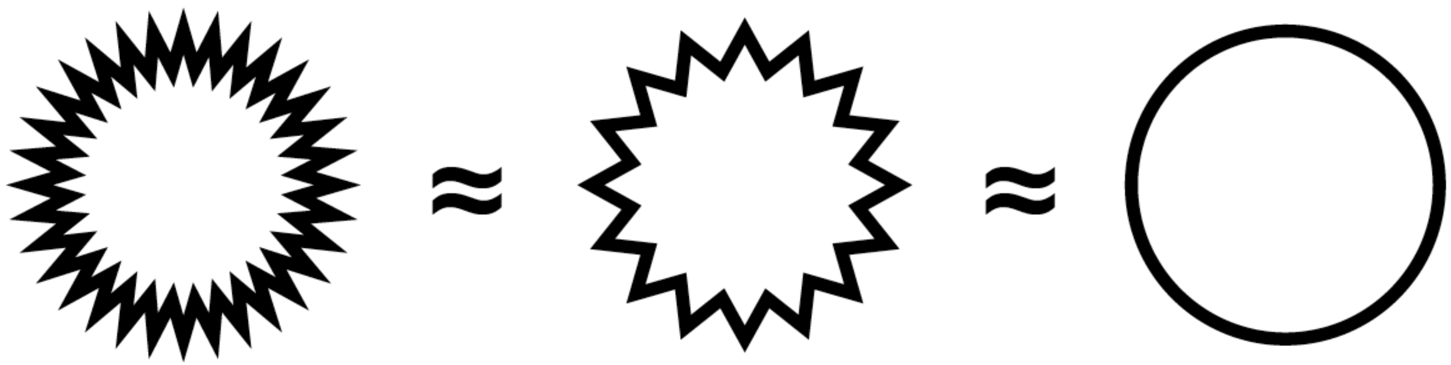
From < 1s to a few hours / event.

- ~ 50 MC generators on the market. *How many can you name?*
- >> 50 combinations of MC generators in a sample.
- ~ 35 K samples generated on ATLAS in the last "campaign" of 2012.
- ~ 7 B events!

Courtesy: Z. Marshall

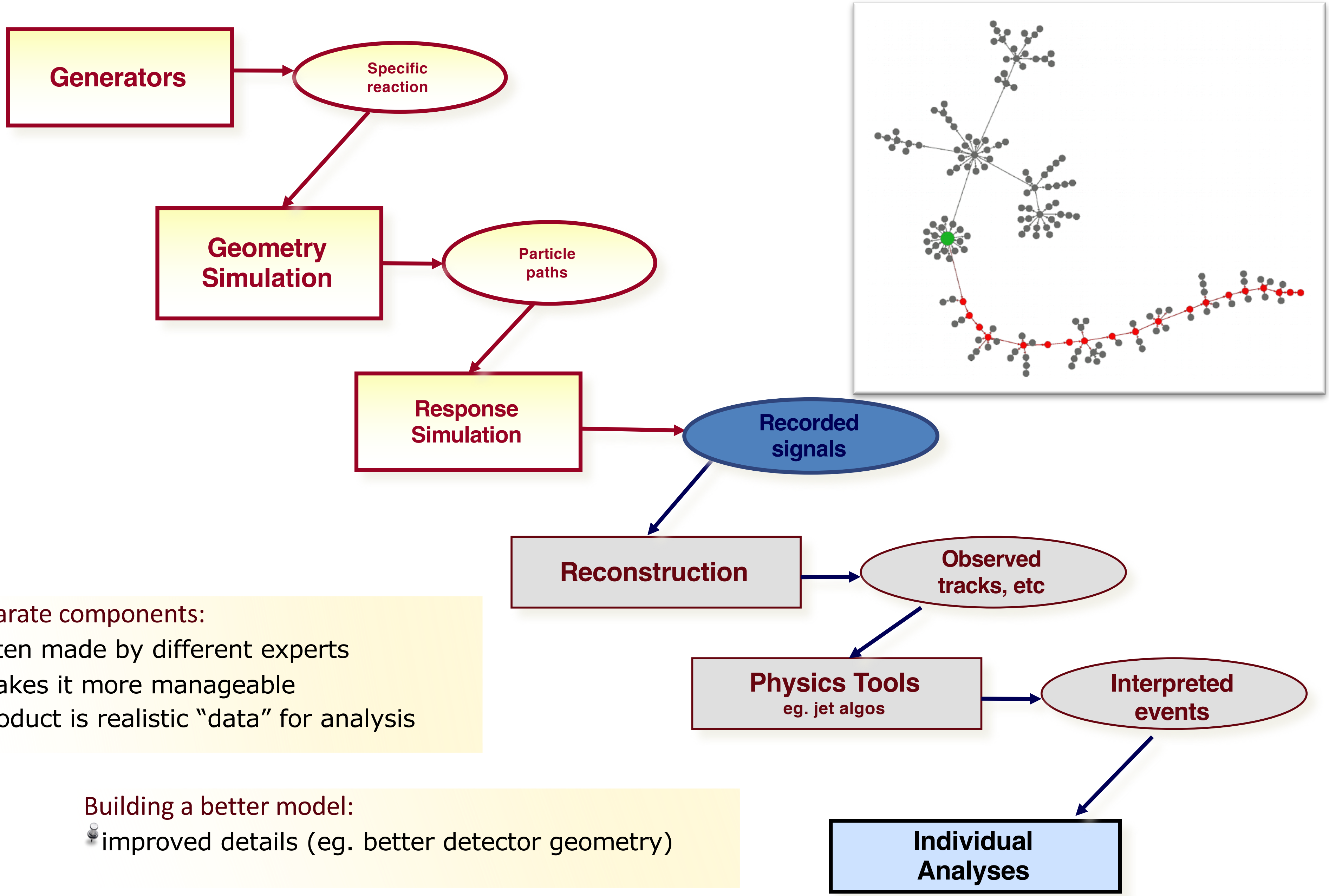
SIMULATION – HOW

1. **Break the problem up as much as possible.**
   - Do you understand all the steps of the system?
2. **For each piece of the problem, write some code**
   - Did you remember all the effect for each step?
3. **Figure out what accuracy is needed.**
   - And spend the appropriate time in working out the details.



4. **Cross your fingers and press the button.**

Flow of simulated data



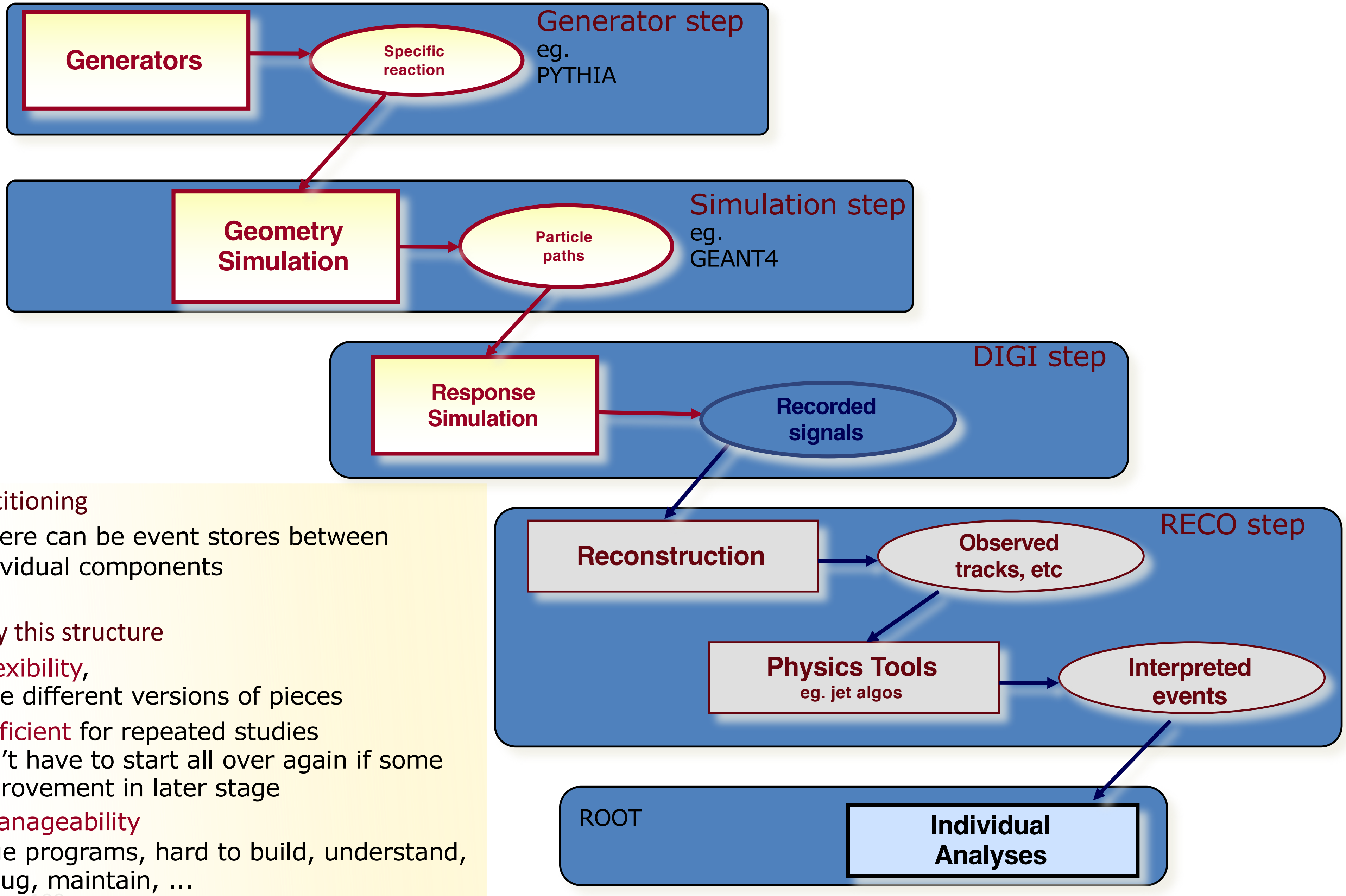
Separate components:

- often made by different experts
- makes it more manageable
- Product is realistic "data" for analysis

Building a better model:

- improved details (eg. better detector geometry)

Partitioning production systems



Partitioning

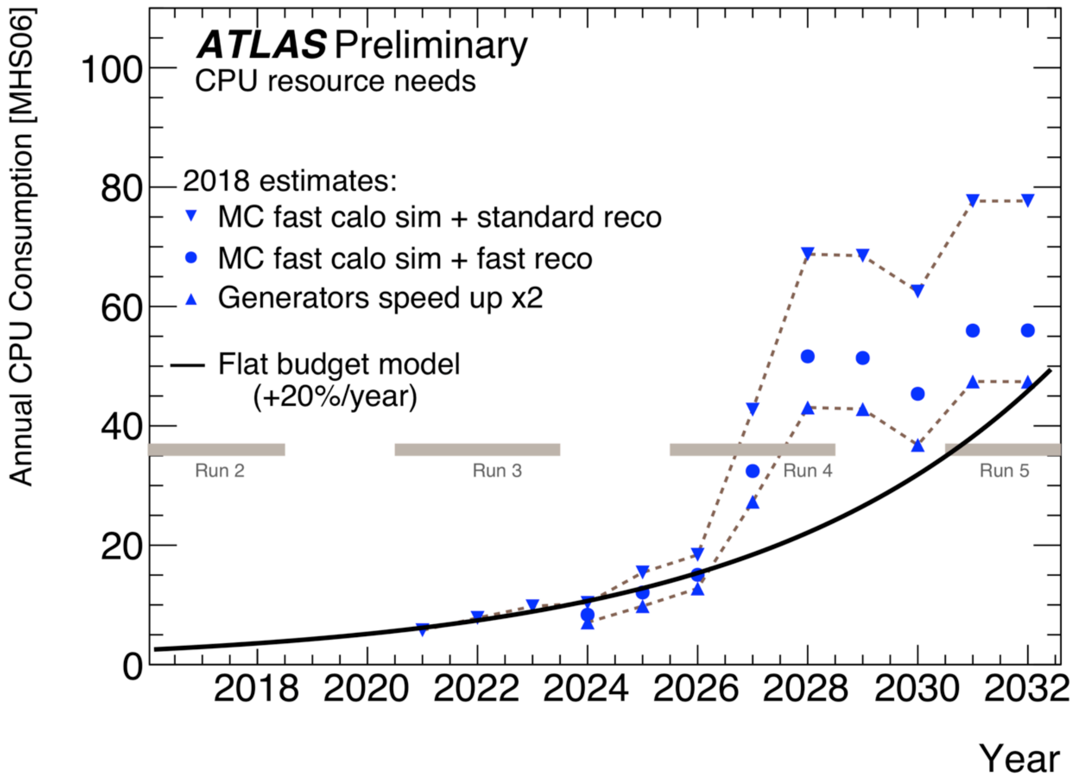
- there can be event stores between individual components

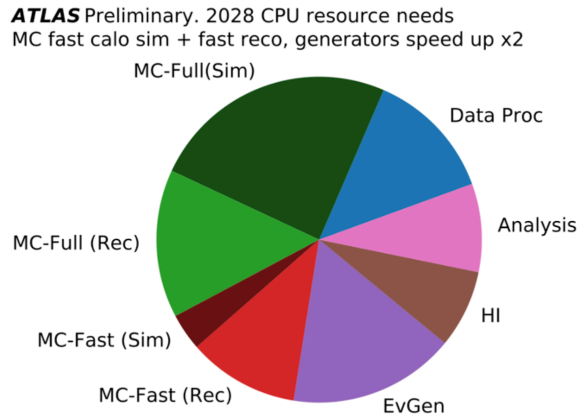
Why this structure

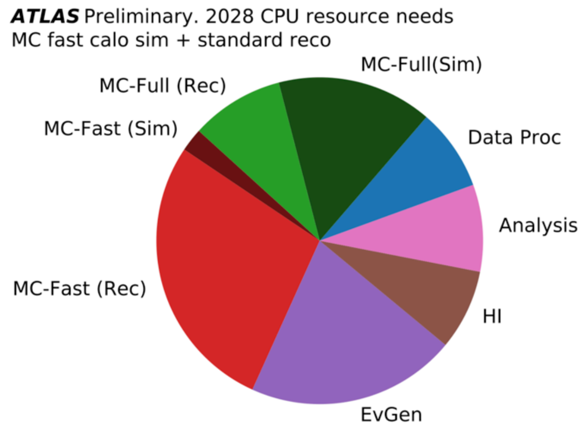
- flexibility,
have different versions of pieces
- efficient for repeated studies
don't have to start all over again if some improvement in later stage
- Manageability
large programs, hard to build, understand, debug, maintain, ...

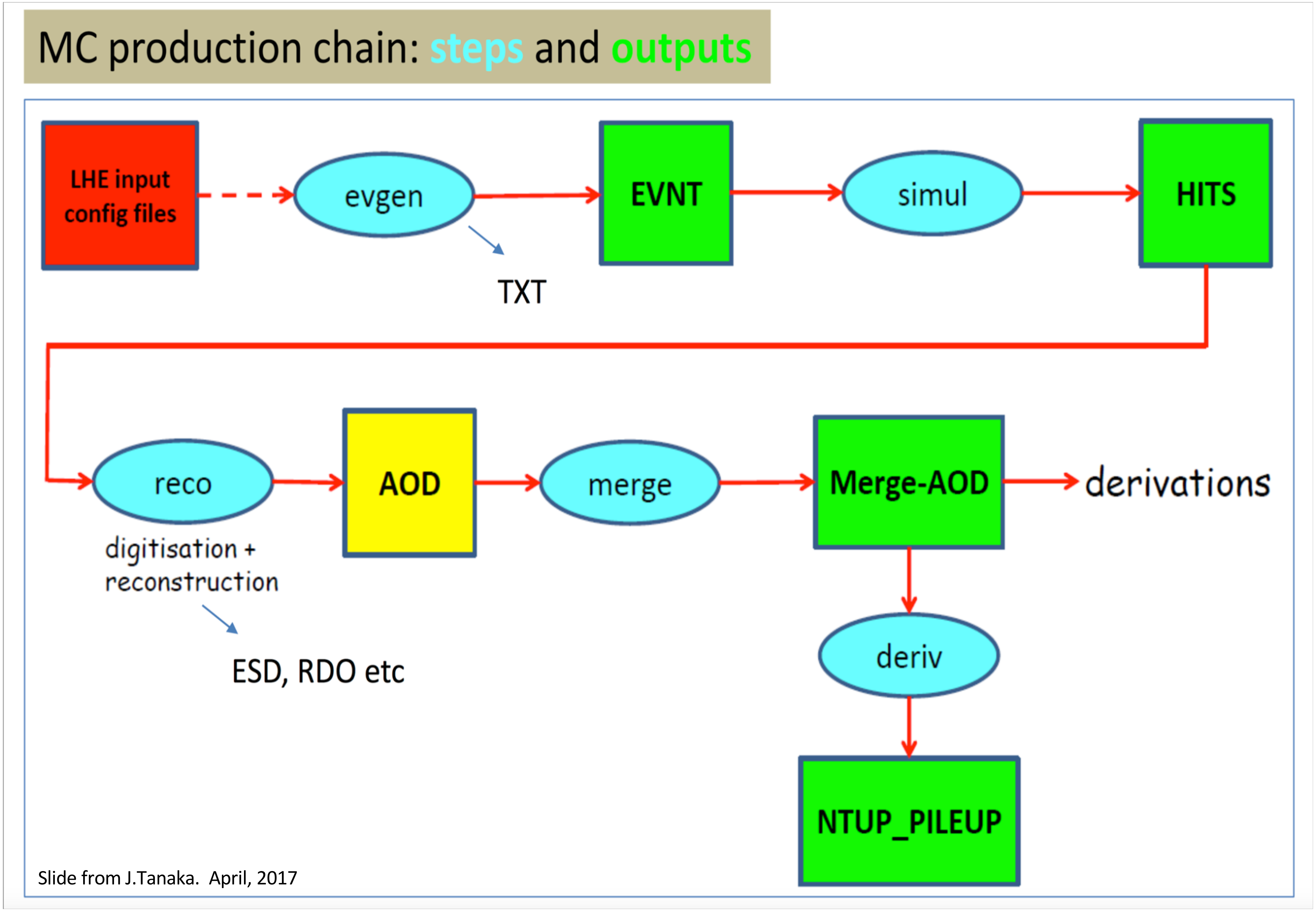
Simulation needs for HENP scientific program ?

Reason 1: we will need them to be able to exploit the HL-LHC



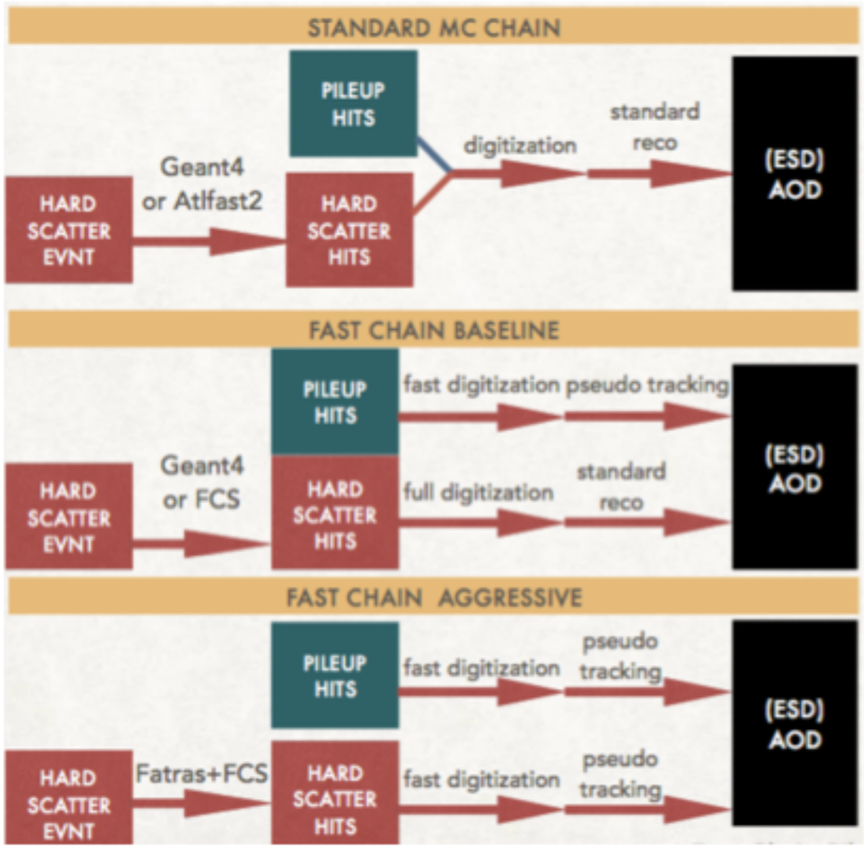




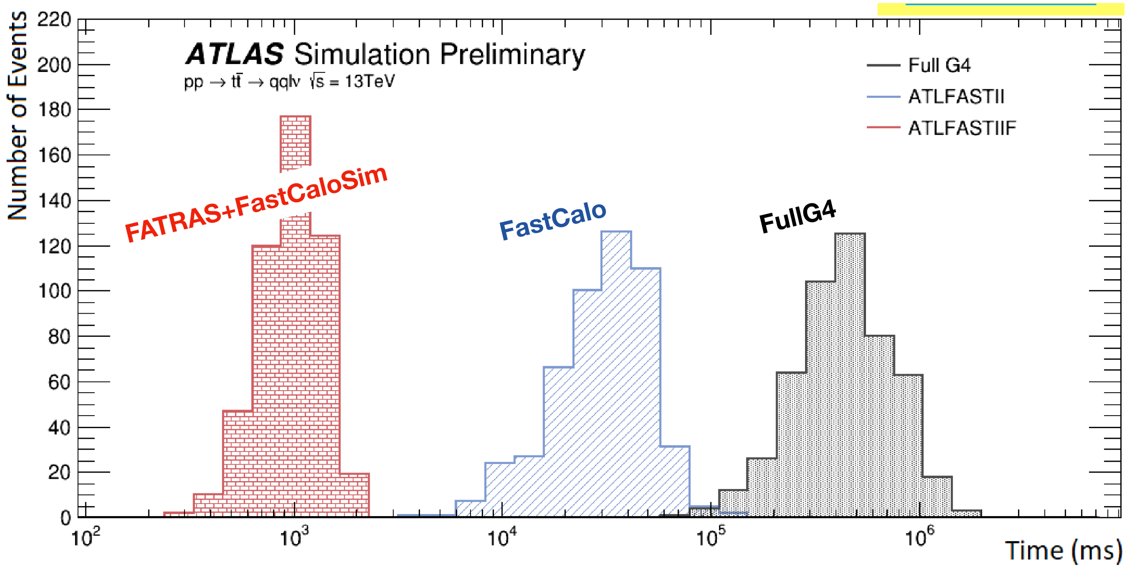

MC production chain: steps and outputs
LHE input config files
evgen
TXT
EVNT
simul
HITS
reco
digitisation + reconstruction
ESD, RDO etc
AOD
merge
Merge-AOD
derivations
deriv
NTUP_PILEUP
Slide from J.Tanaka. April, 2017

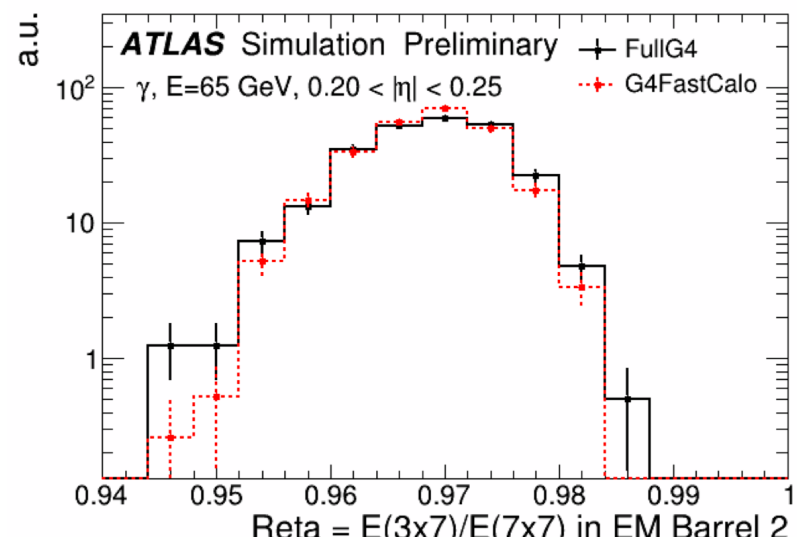
Fast Simulation Chain

- To be used to speed up of tracker simulation chain
- Can use full MC chain for hard scatter and fast chain simulation of hadron-hadron for pileup : more detailed treatment for the important part of event
- Components :
  - Simulation (FATRAS)
  - Silicon digitization
  - TRT digitization
  - Pseudo-tracking : seed tracks with PrepRawData associated to truth track, to skip the CPU consumption track finding steps
- At least 1/4 of CPU to be used for Full simulation
  - Tuning and improvement of full simulation very important. Already underway
- Fast Chain is a key ingredient for future optimizations
  - Validation of physics results is a major challenge
- New ideas are needed, e.g. GANs



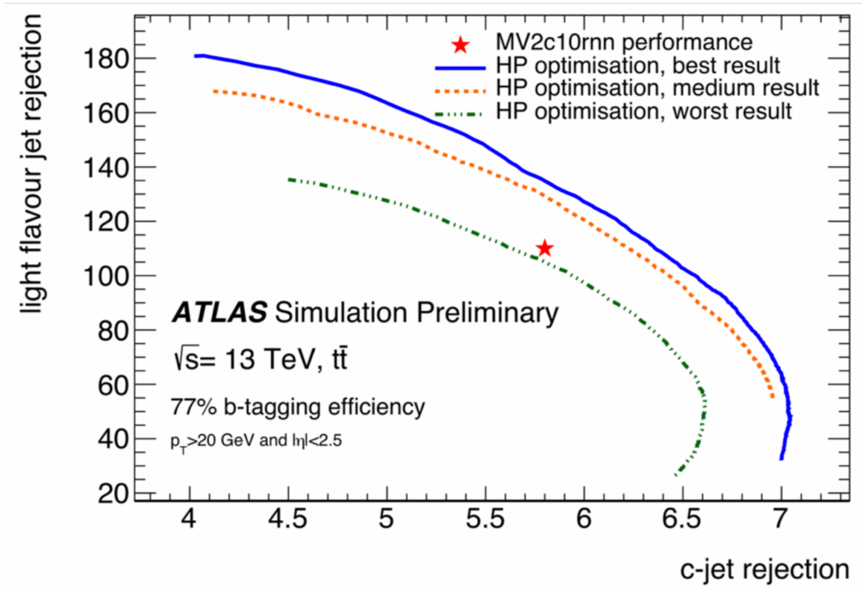
FATRAS about 400 times faster than full G4

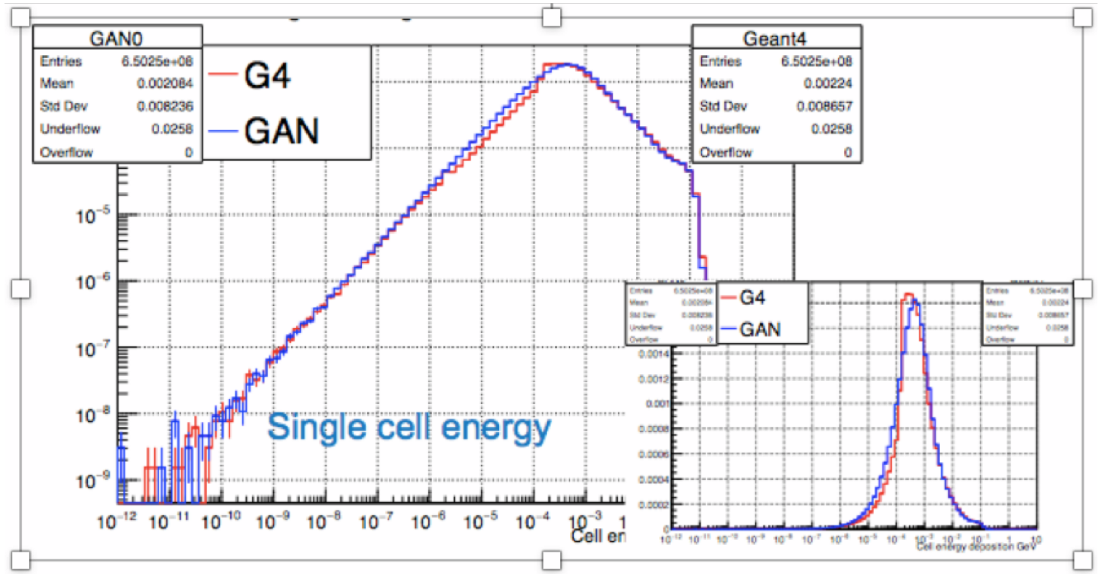




Machine Learning

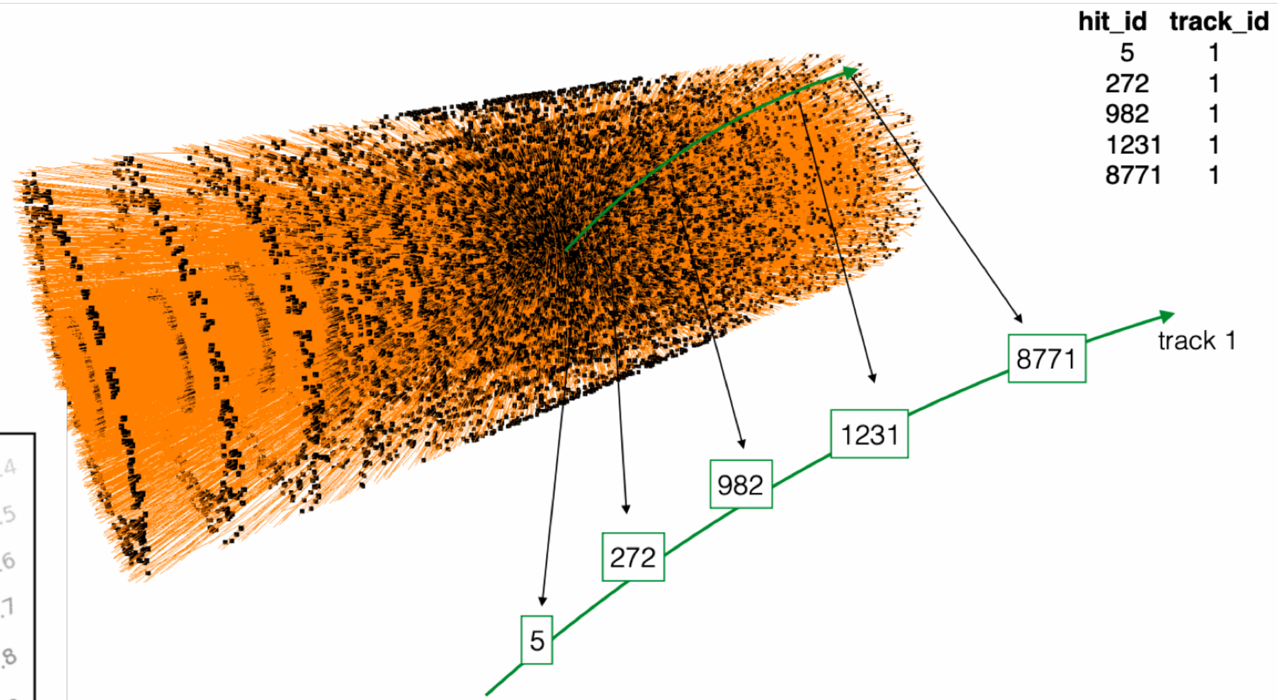
- Machine Learning is a fashion at the moment :
  - ML assisted reconstruction used by HEP experiments
  - New generation of HEP physicists expect to use it (career progression)
  - It could radically speed up simulation and reconstruction at the inference step
- Current usage still at the ML-driven reconstruction and analysis
- We are still better than a machine
  - HEP has a long tradition of data analysis (which today is called data science)
- Started to look how to integrate industry standard ML systems into our software framework

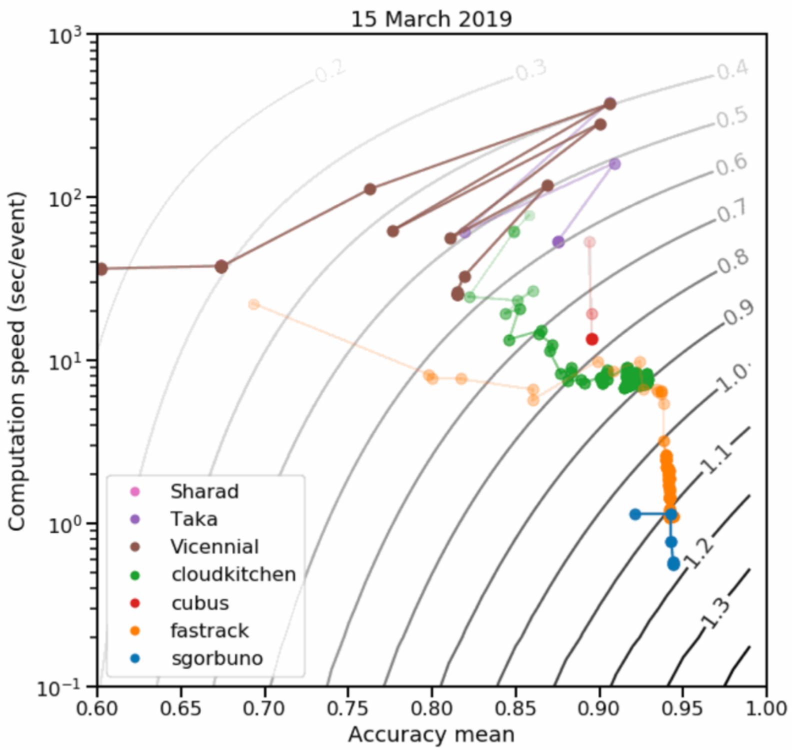




Tracking and ML

- Tracking is a major CPU consumption in reconstruction
- trackML challenge



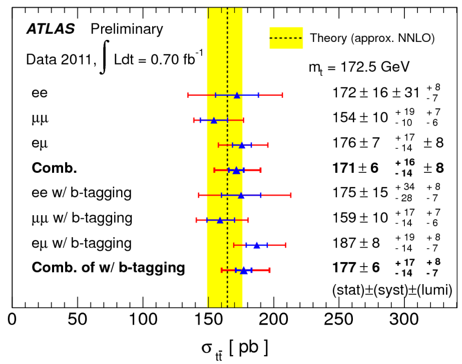


Physics Analysis Steps

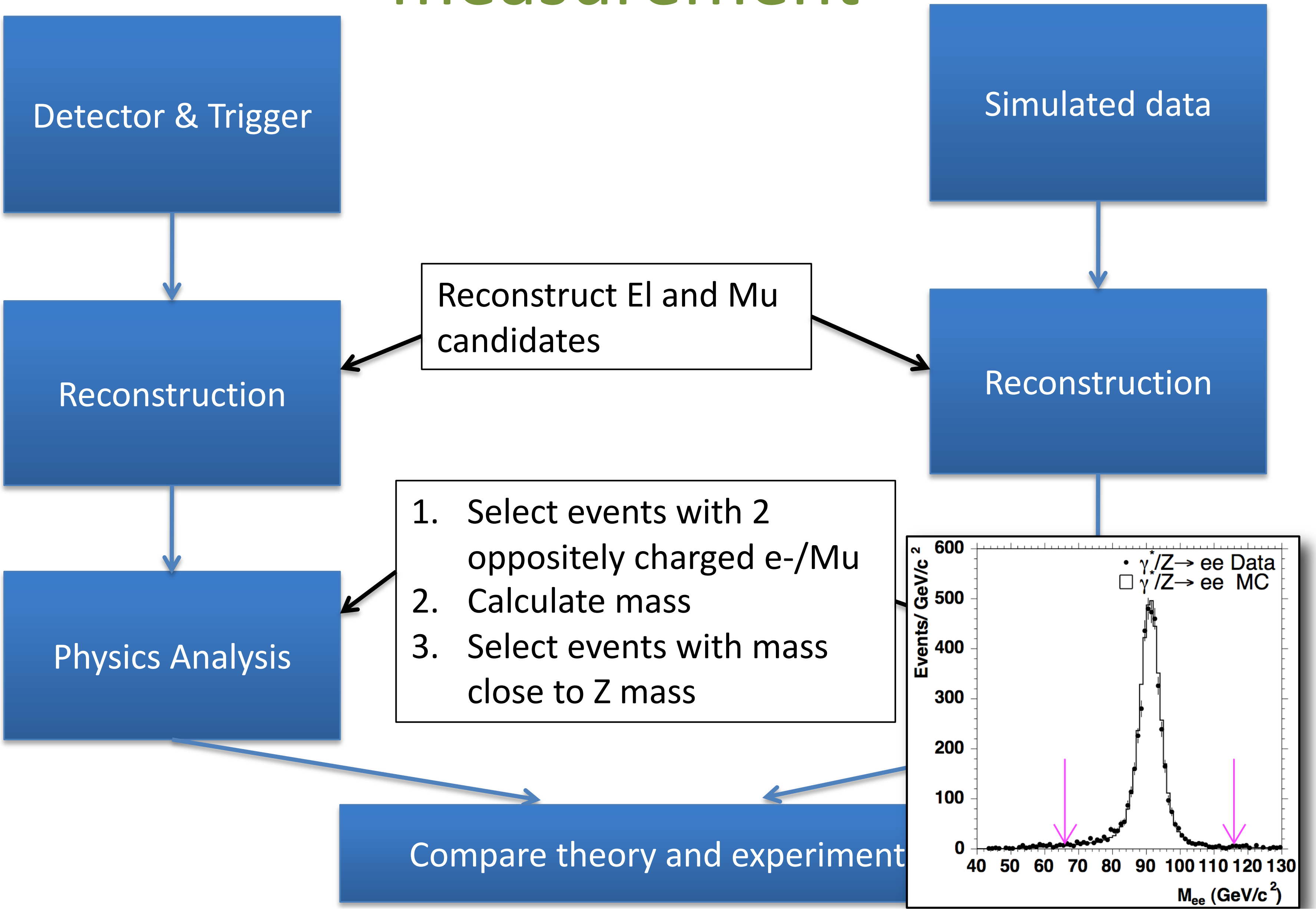
~TB

- Start with the output of reconstruction
- Apply an event selection based on the reconstructed object quantities
  - Often calculate new information e.g masses of combinations of particles
  - Event selection designed to improve the ‘signal’ to ‘background’ in your event sample
- Estimate
  - Efficiency of selection (& uncertainty)
  - Background after selection (& uncertainty)
  - Can use simulation for these – but have to use data-driven techniques to understand the uncertainties
- Make final plot
  - Comparing data to theory
  - Correcting for efficiency and background in data
  - Include the statistical and systematic uncertainties

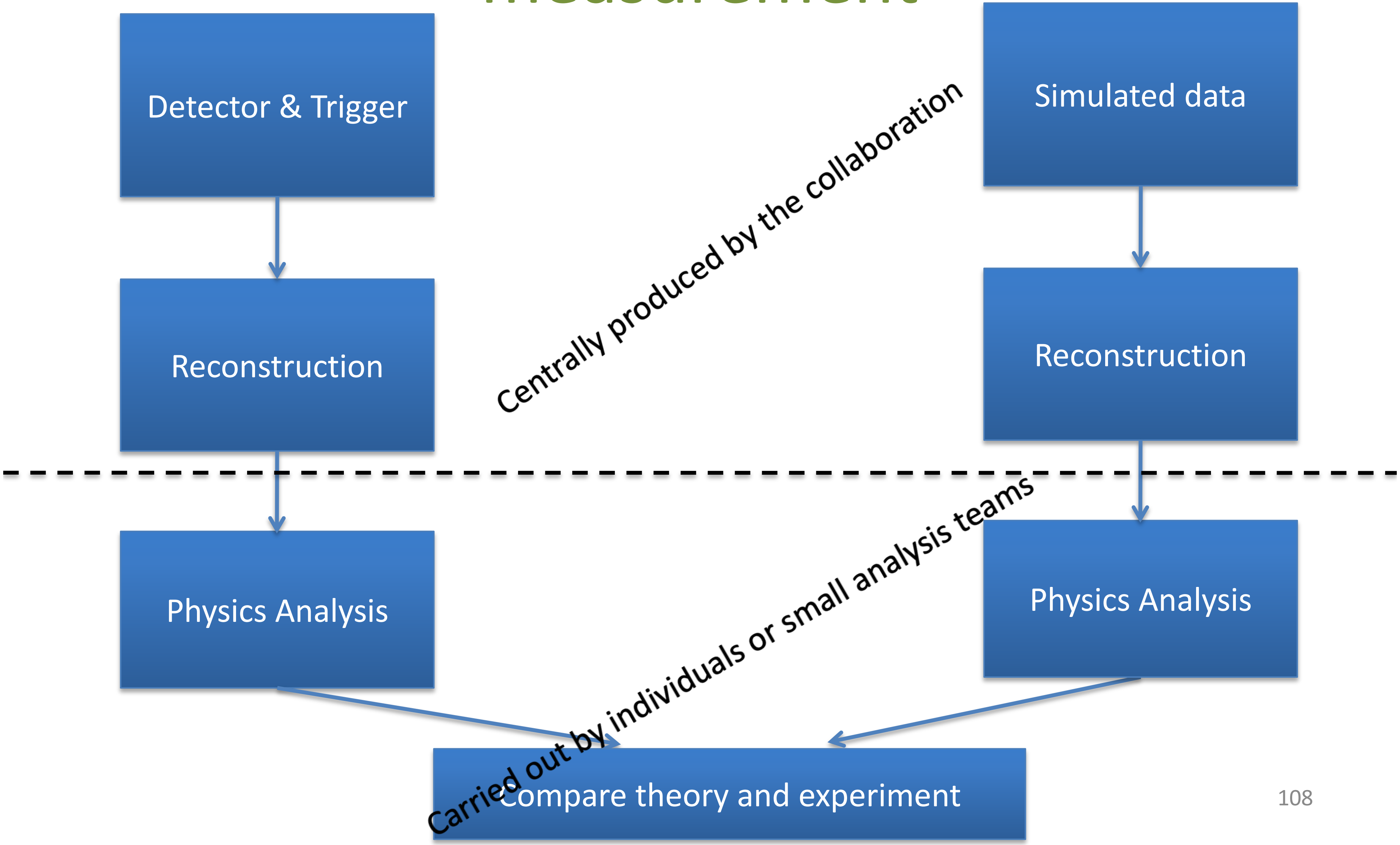
~kB



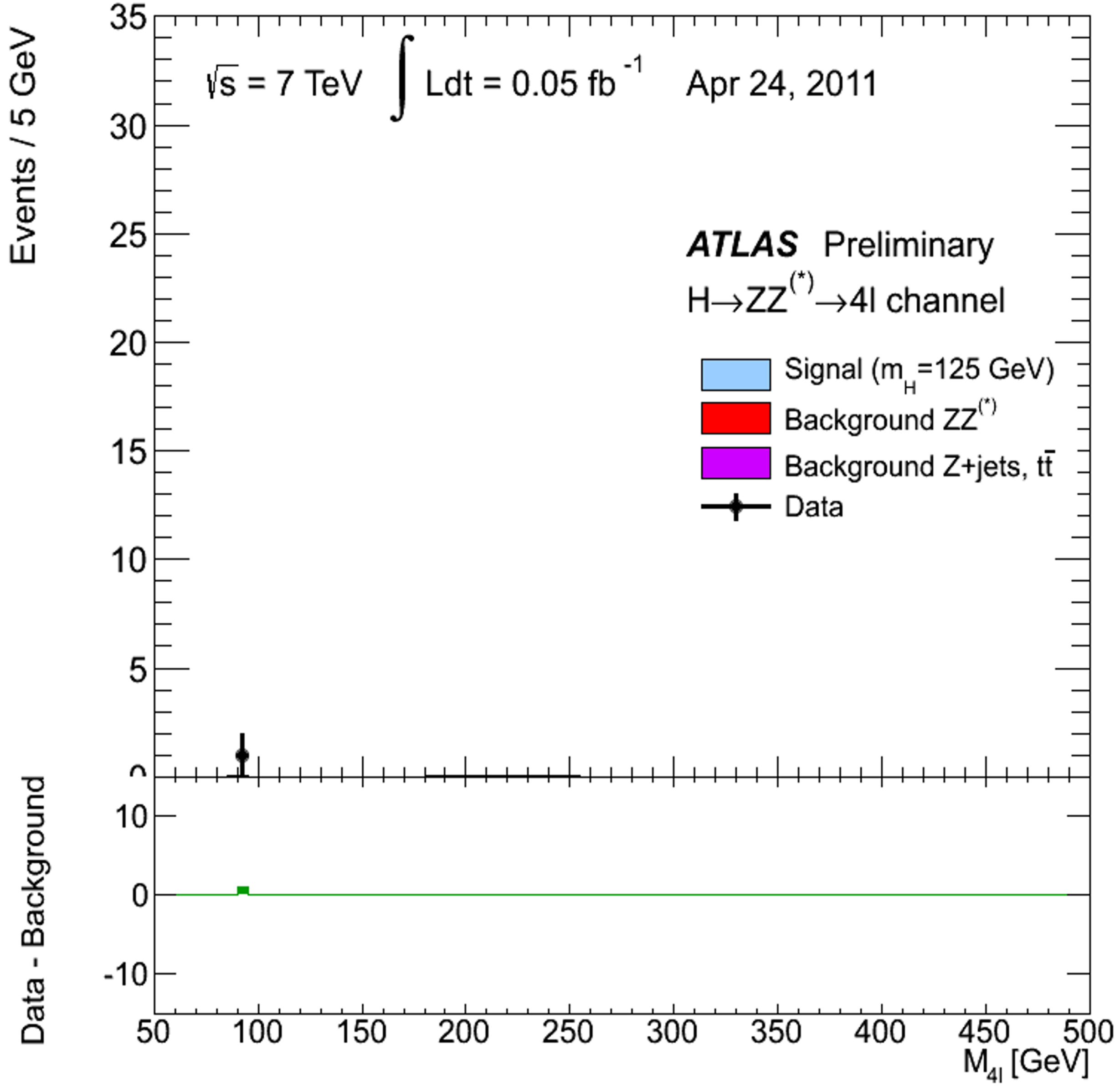
Analysis flow in Z cross-section measurement



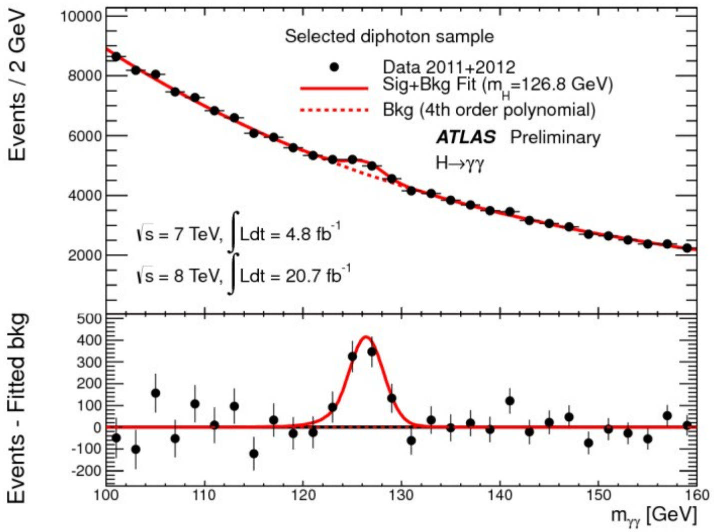
Analysis flow in Z cross-section measurement

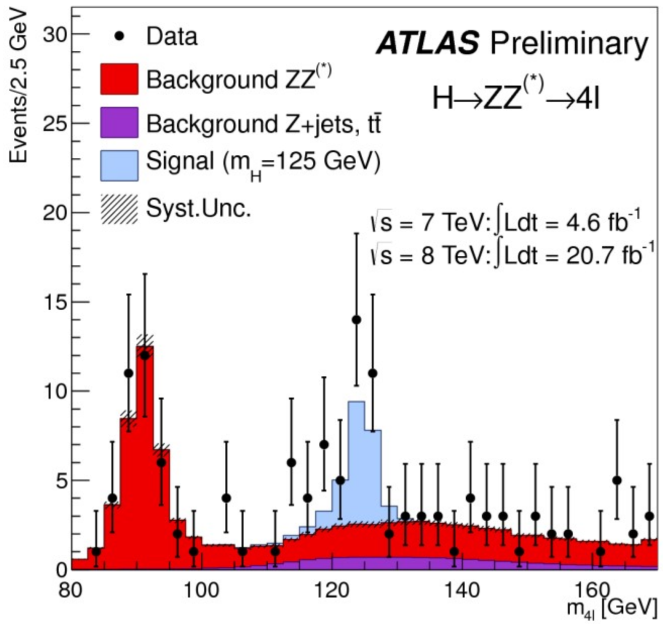


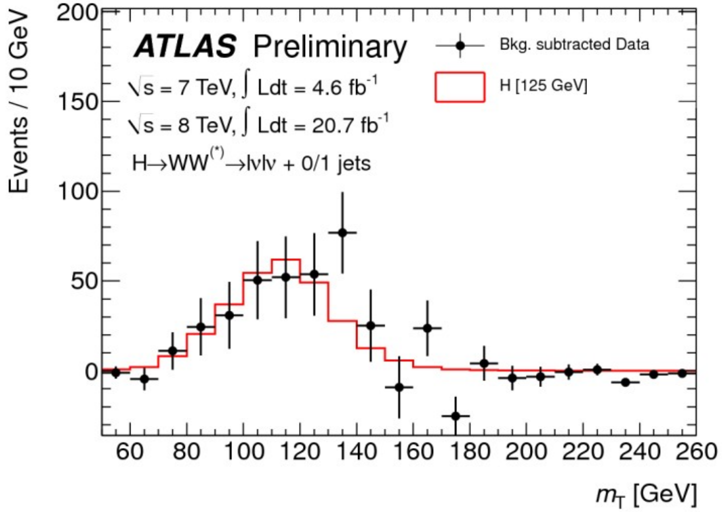
Higgs Boson Discovery



Evidence for a new Particle



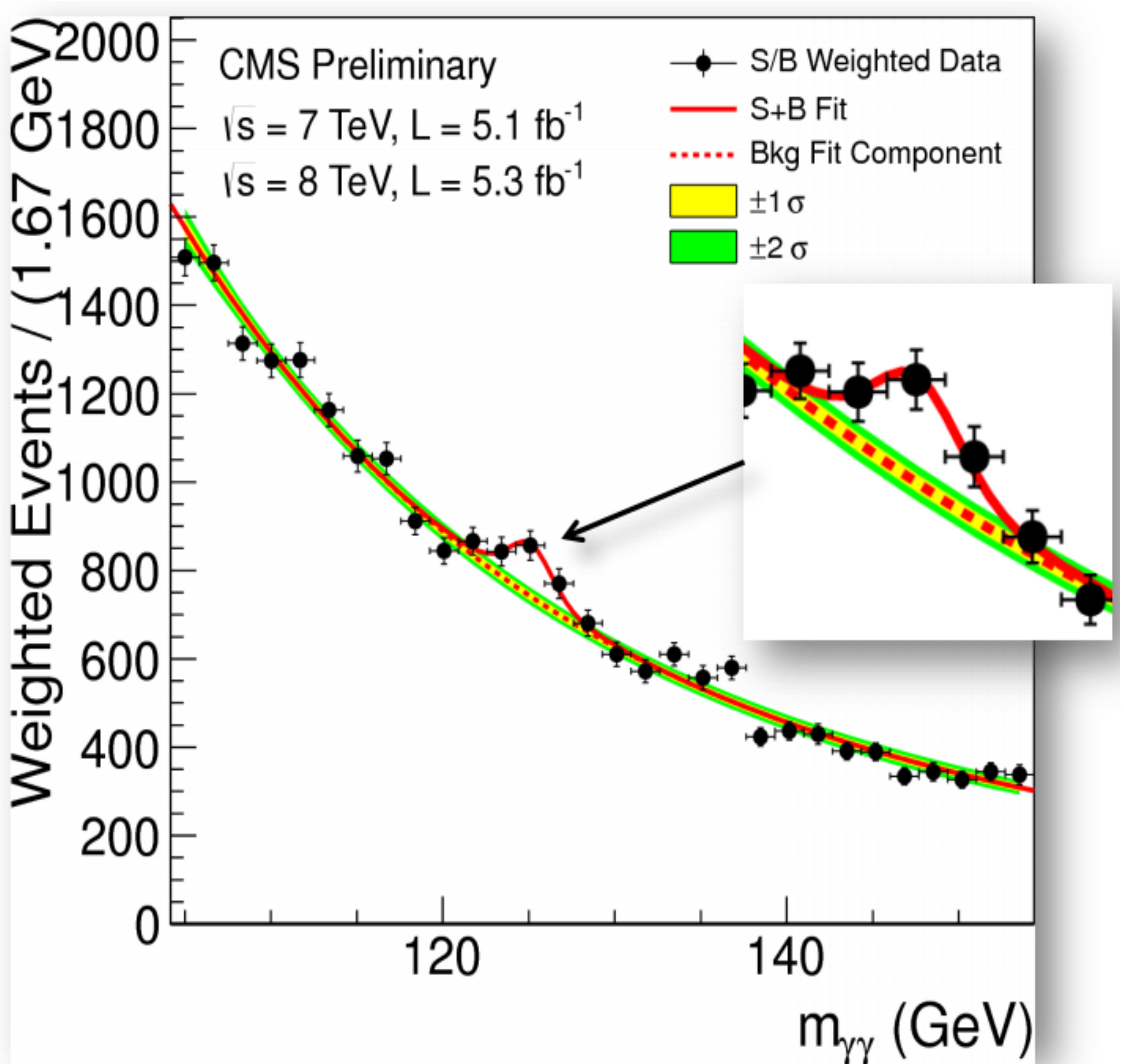




- γγ, ZZ and WW channels updated with full data

- Clear evidence for a new particle seen in these channels

H -> γγ Results



Analysis carried out in different categories of events with expected different resolution and Signal /Background.

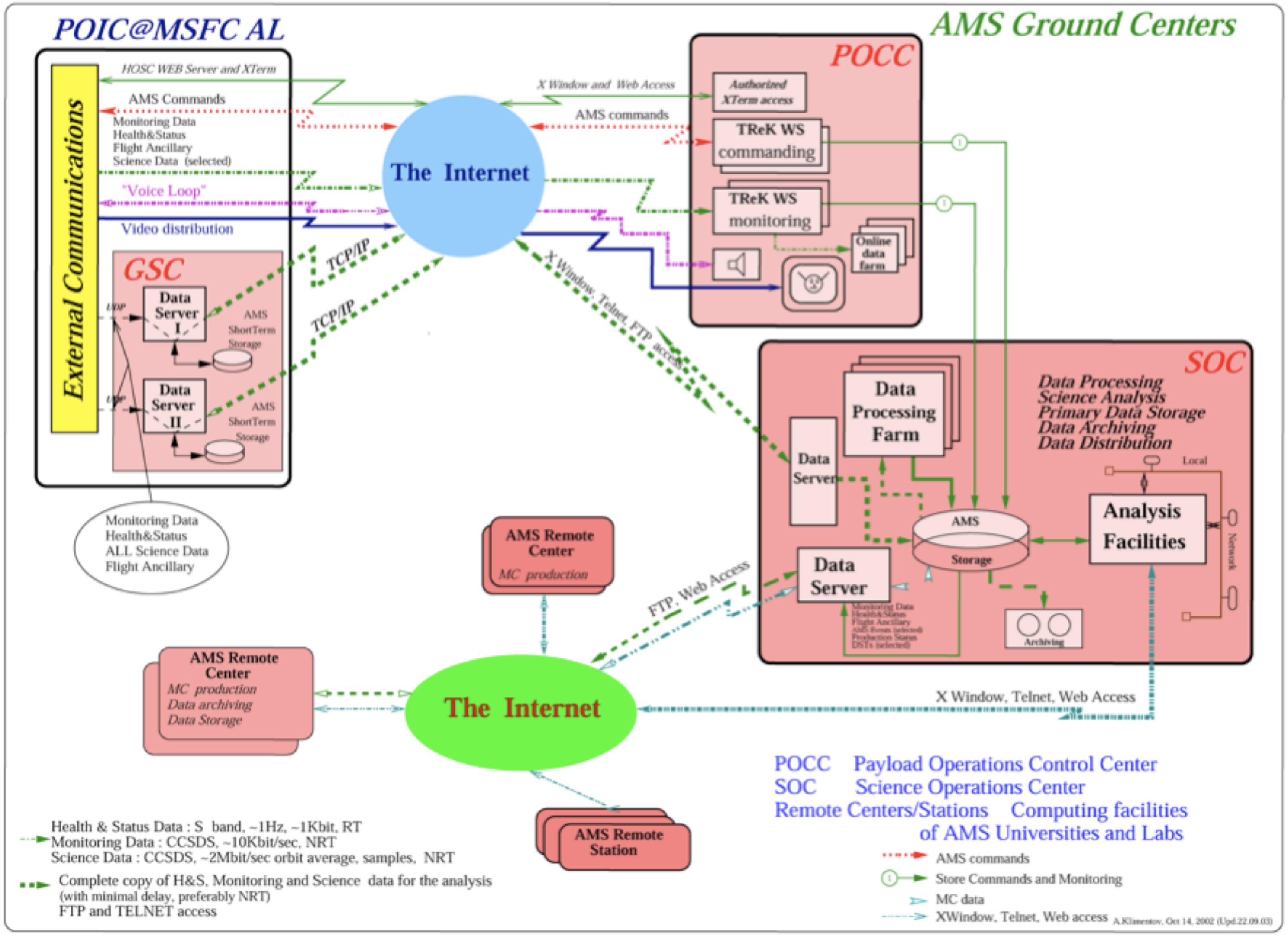
Left plot shows mass plot where events are WEIGHTED by the category (more like what the fit "see's").

ATLAS & CMS both saw significant (>4σ) peak at ~126GeV. Most important result in the recent Higgs observation.

Offline Computing in High Energy Physics

- Offline computing in HEP
  - Has changed and evolved dramatically over the past decades
  - Especially for the biggest experiments – at the LHC
- The situation ~15 years ago
  - Data processing was performed in large computing centers using local batch systems with dedicated shares
  - A few satellite centers did simulations, occasionally data reprocessing
  - Users were mostly located near large computing centers, usually at the laboratory where the experiment was located, and used a combination of desktops and batch systems for analysis
  - Final Data Summary Files versions were physically shipped for remote analysis

Компьютерная модель эксперимента AMS-02 на Международной Космической Станции



AMS - Alpha Magnetic Spectrometer [AMS-01: shuttle flight @1998, AMS-02 : ISS 2010-now]

Распределенная модель компьютинга для экспериментов на БАК. Экономические и социологические причины :

- Объем данных БАК, не позволял просто расширить существующий ВЦ ЦЕРН и использовать только один ВЦ для хранения, обработки и анализа данных.
  - Требовались капитальные вложения в инфраструктуру и в случае использования централизованной модели взнос стран участниц в бюджет организации мог возрасти в несколько раз, при этом ЦЕРН должен был одновременно обеспечить строительство самой “машины” и сопутствующей инфраструктуры;
- Количество ученых, участвующих в экспериментах на БАК, уже на первом этапе заявок было близко к 5 тысячам (в настоящее время около 9 тысяч) из более чем 50 стран мира, в случае централизованного решения, анализ данных в ЦЕРН создавал неравноправные условия для стран находящихся на значительном расстоянии от ЦЕРН (таких как Россия, США, Япония, Австралия, Канада), доступ к данным для них был бы не столь эффективен как для стран западной Европы;
- Многие страны, Университеты, Исследовательские институты имели значительные вычислительные мощности, и были заинтересованы в их развитии и использовании;
- Экономическая ситуация во многих странах мира требовала вложений в национальные проекты и создания рабочих мест в странах ЕС, поэтому идея дополнительного финансирования компьютерных мощностей ЦЕРН не была поддержана экспериментами. Одновременно идея о расширении национальных ВЦ для потребностей LHC была воспринята позитивно мировым сообществом.

Распределенная Модель Компьютинга для экспериментов на БАК. Технические причины :

- Ни ЦЕРН, ни другие центры ФВЭ и ЯФ не имели опыта строительства ВЦ для обработки данных в мульти-петабайтном диапазоне и одновременного доступа к данным тысяч пользователей.
- Характеристики будущего центра в части потребляемой мощности и систем охлаждения не могли быть реализованы на территории ЦЕРН в Швейцарии или Франции без изменения двухсторонних соглашений организации с этими странами;
- Использование суперкомпьютера (или нескольких СК) для проведения централизованной обработки данных не позволяло решить вопрос с анализом данных, не говоря о стоимости такого решения; ПО физических экспериментов (а это 4М инструкций кода) не было оптимизировано для СК и, в частности, для графических процессоров;
- Технологии иерархического гибридного хранения данных (диск-лента), не позволяли эффективно мигрировать файлы между постоянным (лента) и временным (диск) хранилищами с эффективностью и объемами, требуемыми для обработки будущих данных LHC в одном центре;
- Оценка возможностей WAN не гарантировала эффективный удаленный доступ к данным;
- Требования к вычислительному и дисковому ресурсу значительно менялись в течении подготовки экспериментов на БАК.

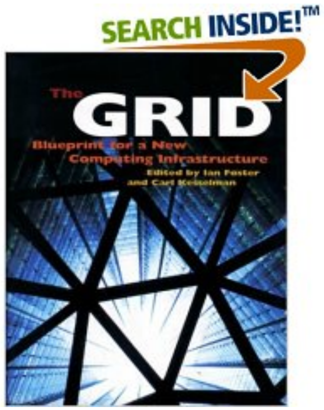
Распределенная Модель Компьютинга для экспериментов на БАК

- Грид технологии были предложены в конце прошлого века Я.Фостером и К.Кессельманом и основная концепция изложена в книге «GRID: a Blueprint to the New Computing Infrastructure»

- Задачи ФВЭ и ЯФ привели к широкому использованию грид технологий.
  - Еще на раннем этапе развития компьютерной модели LHC (конец XX века) было принято решение объединить существующие и вновь создаваемые вычислительные центры (более 300 центров на сегодняшний день) в распределенный центр обработки данных
  - Для такого решения были причины технического, экономического и социологического характера.

WEB and Grid

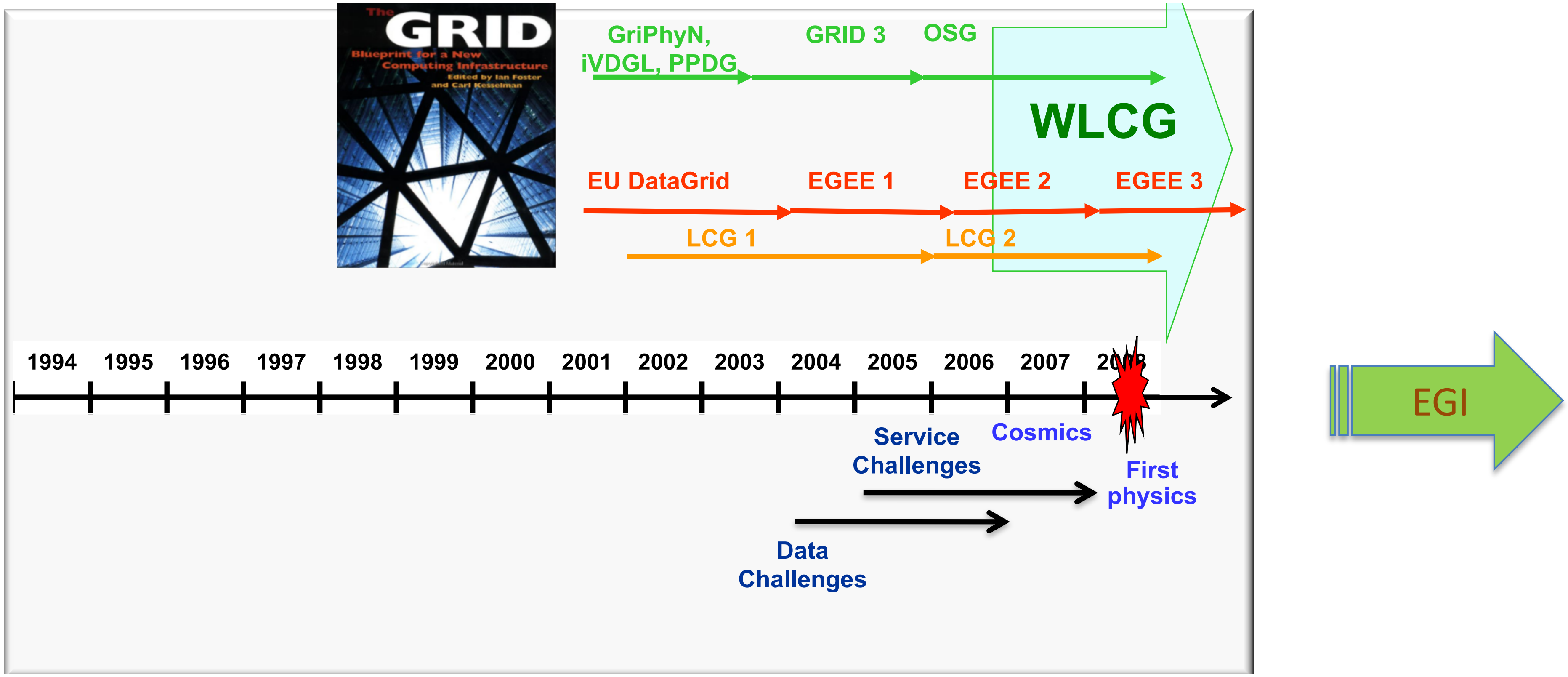
The World Wide Web provides seamless access to information
that is stored in many millions of different geographical locations



The Grid is an infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe

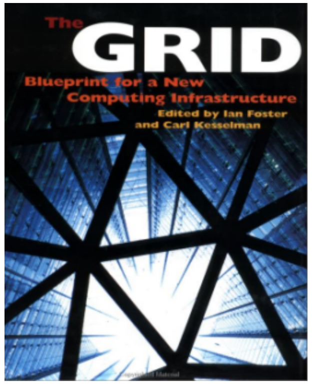
Tim Berners-Lee invented the World Wide Web at CERN in 1989

Foster and Kesselman 1997



- When LHC computing was started (~2001)
  - There were no internet companies, no cloud computing – Google was a search engine, Amazon, etc. did not exist
- We had to invent all of the tools from scratch
  - At CERN we had no tools to manage a data centre at the scale we thought was needed (no commercial or OS tools existed)
    - Initial tools developed through EU Data Grid
- Grid ideas from computer science did not work in the real world at any reasonable scale
  - We (EU, US, LHC grid projects) had to make them work at scale
  - We had to invent trust networks to convince funding agencies to open their resources to federated users
- Physics community was not convinced that any of this was needed ;-)
- Grid initiative in Russia – RDIG – Russian Data Intensive Grid (V.A.Ilyin)

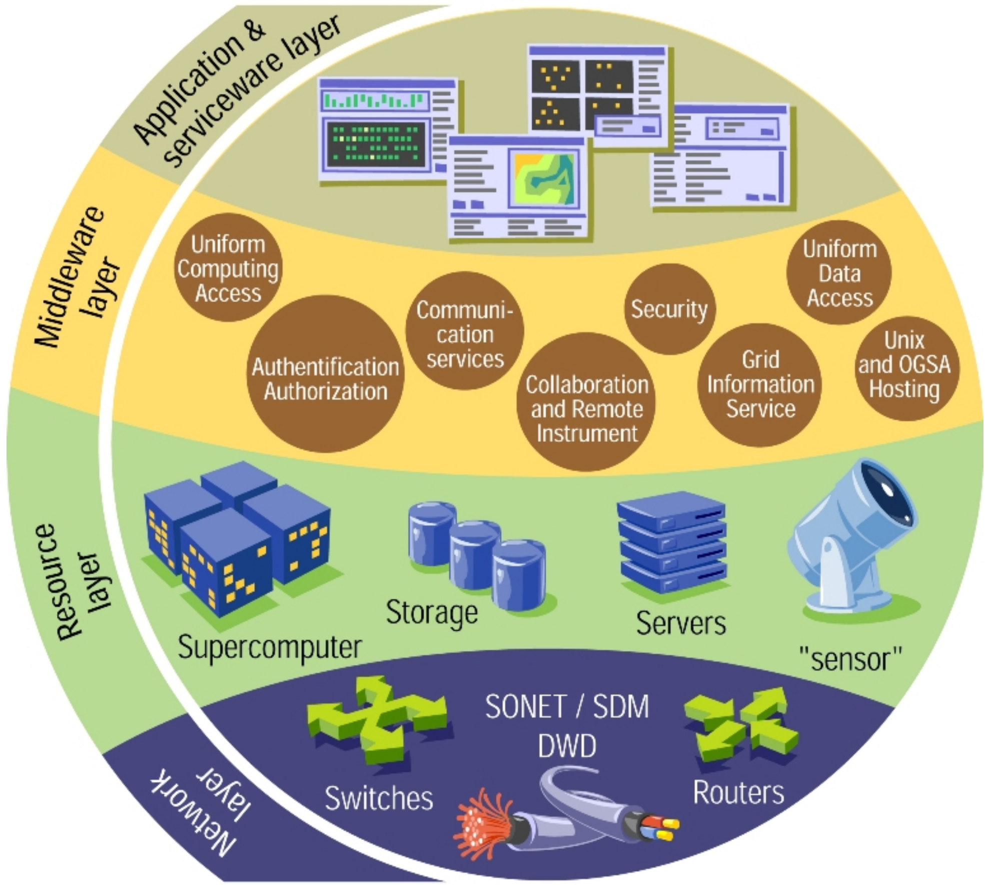
What is a computing Grid ?



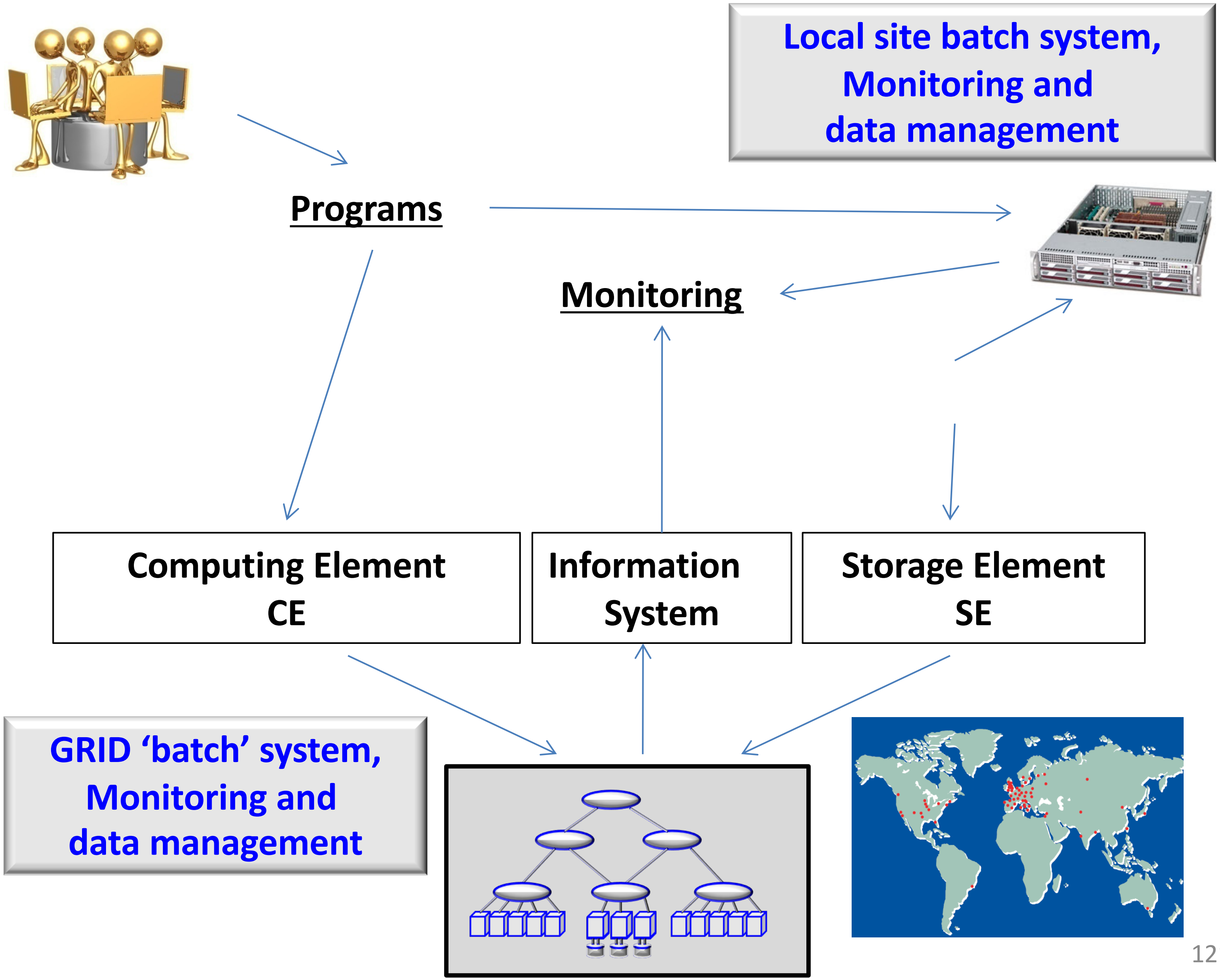
- There are many conflicting definitions……
- **<u>1998 The Grid by Ian Foster and Karl Kesselman</u>**
  - Made the idea popular
- "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations. "
  - These are the people who started globus, the first grid middleware project
- From the user's perspective:
  - I want to be able to use computing resources as I need them
    - I don't care who owns resources, or where they are
    - Have to be secure
    - My programs have to run there
- The owners of computing resources (CPU cycles, storage, bandwidth)
  - My resources can be used by any authorized person (not for free)
    - Authorization is not tied to my administrative organization
- **NO centralized control of resources or users**

How Does the Grid Work ?

- It relies on advanced software, called **middleware**.
- Middleware automatically finds the **data** the scientist needs, and the **computing power** to analyse it.
- Middleware balances the load on different resources. It also handles **security, accounting, monitoring** and much more.



How Does Grid Work ?



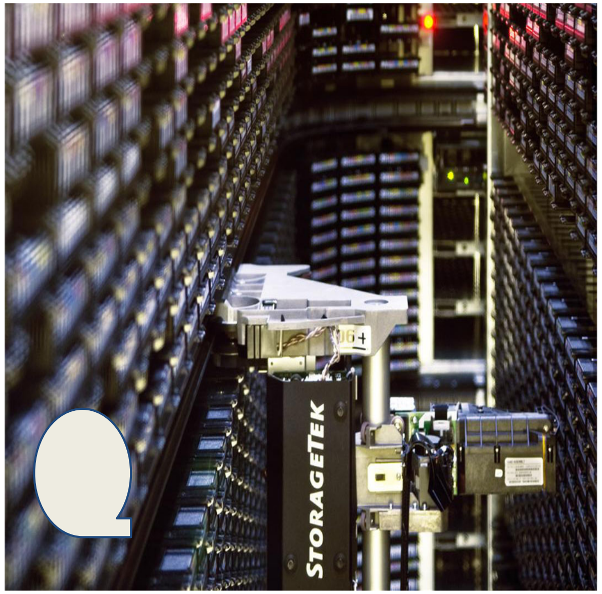
Main components of a WLCG site

CPU servers:

- CPU servers are grouped into **Batch systems** for processing of data

Disk servers:

- CPU server attached to several disks
- Disk servers are grouped into **Storage Elements**
- Data storage with fast access



Tape robots:

- Long term archival
- Slow access
- All raw data from the experiments has 2 tape copies

Grid middleware: the glue

- Heterogeneous resources are grouped and exposed in a uniform way
  - Computing Elements give access to CPUs
  - Storage Elements give access to data
  - Information systems describe the grid
  - Authentication is done via x509 public key infrastructure

Why not store and process everything at CERN?

- Traditionally a single computing center at CERN could not physically provide all resources
- Data redundancy
  - LHC and ATLAS operation is expensive: we can't afford to lose any data
  - There are multiple copies, in particular of RAW detector data
- Funding reasons
  - ATLAS is an international collaborations with participants from 38 countries
  - Funding agencies prefer to invest and employ locally

http://ssl-computing.web.cern.ch/ssl-computing/default.htm

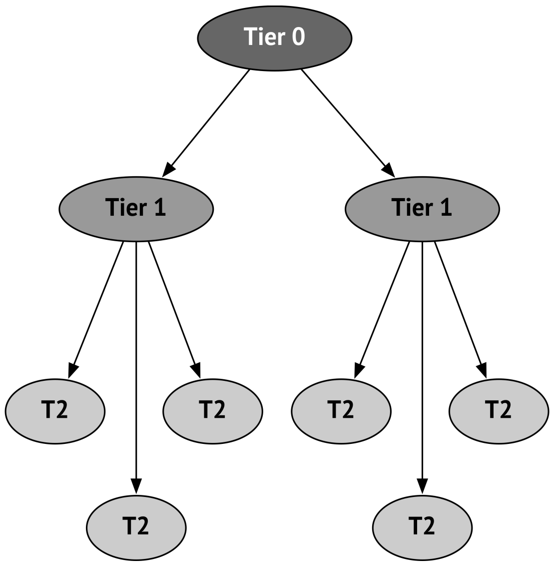
LHC Computing Grid

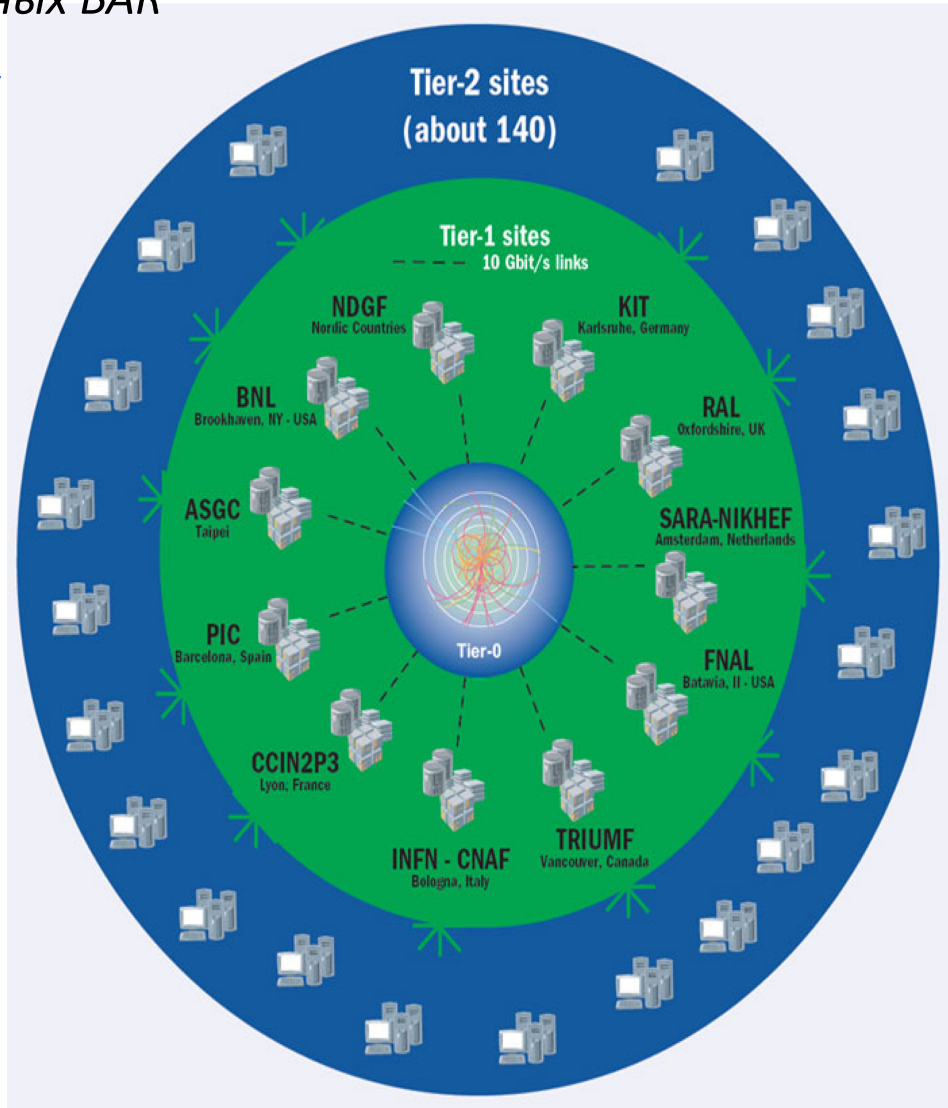
- Collaborating computing centres
- Interconnected with good networking
- Interfaces and protocols that enable the centres to advertise their resources and exchange data and work units
- Layers of software that hide all the complexity from the user
- So the end-user does not need to know where his data sits and where his jobs run
- The Grid does not itself impose a hierarchy or centralisation of services
- Application groups define **Virtual Organisations** that map users to subsets of the resources attached to the Grid

WLCG – World LHC Computing Grid Консорциум вычислительных центров для обработки и анализа данных БАК

Реализация концепции грид для БАК
EGI/EGEE – European Grid Infrastructure, большинство европейских стран, Россия, Индия, Китай, Япония, Тайвань, Корея
OSG – Open Science Grid, США
NorduGrid – Nordic Grid, скандинавские страны, Швейцария, Словения, Венгрия

MONARC - Models of Networked Analysis at Regional Centres for LHC Experiments.





Tier-0 (ЦЕРН) 15%:

- Хранение данных
- Первичная обработка
- Передача данных в T1

Tier-1 (11 центров) 40%:

- Хранение данных
- Переобработка данных
- Централизованный анализ данных

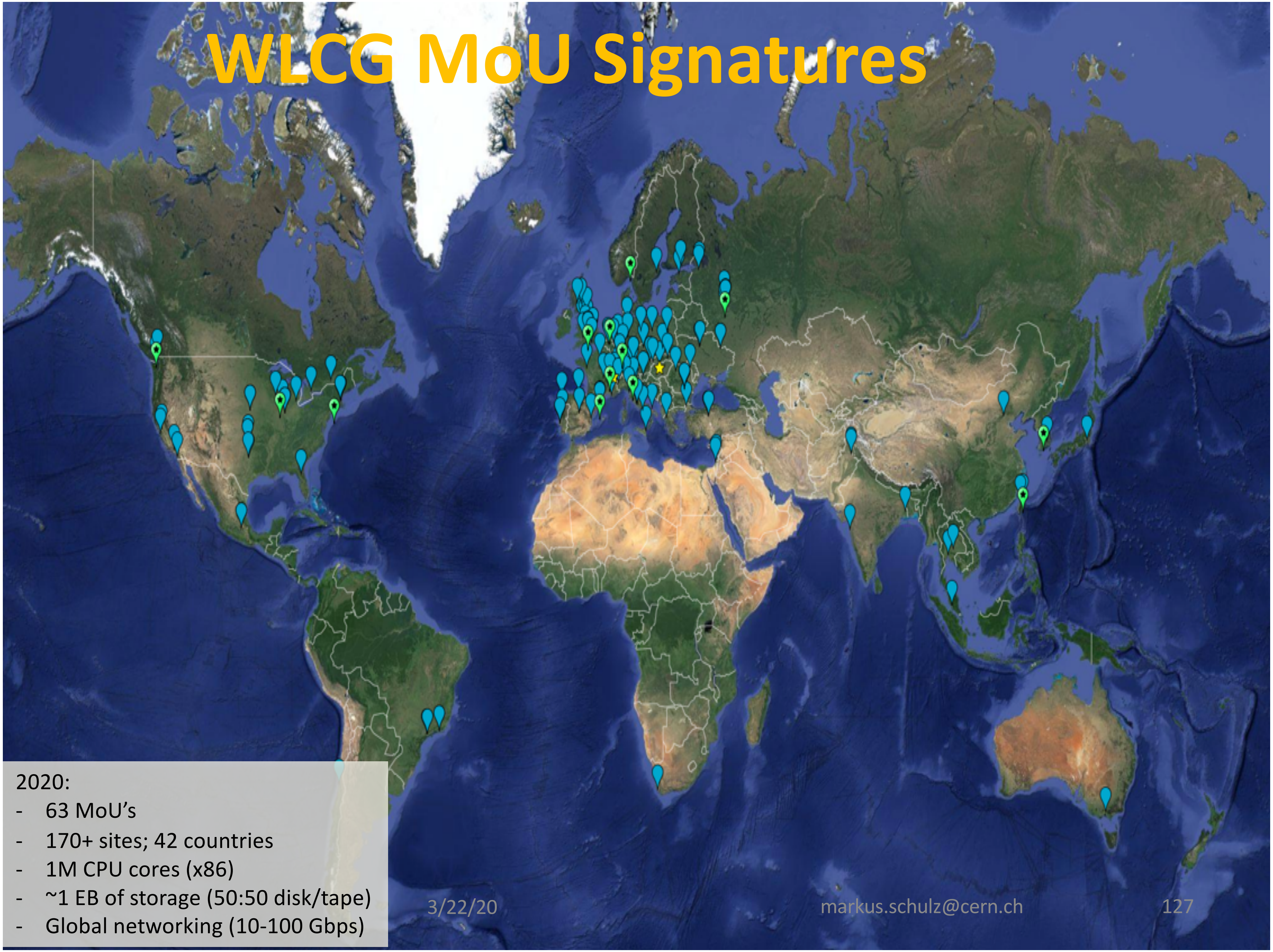
Tier-2 (> 200 центров) 45%:

- Монте-Карло моделирование
- Анализ данных пользователями

Цель создания WLCG : глобальная интеграция вычислительных Центров по всему миру и доступ к их ресурсам для всех участников научной программы LHC

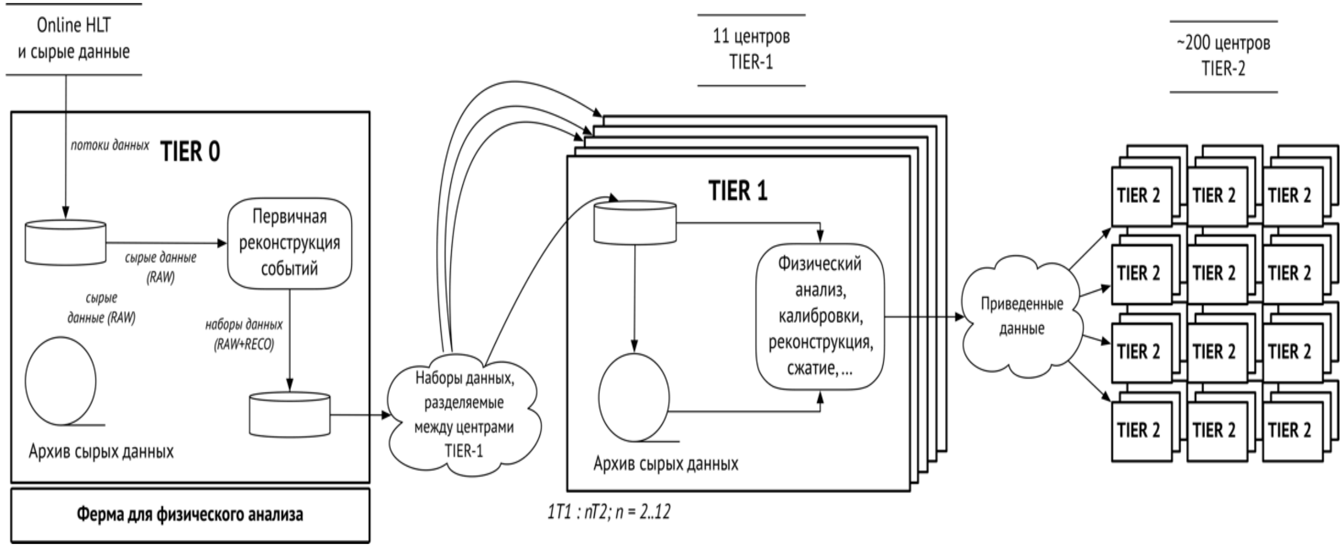
Организация и роль центров консорциума WLCG на момент запуска БАК. 2009 год

WLCG MoU Signatures



2020:
- 63 MoU's
- 170+ sites; 42 countries
- 1M CPU cores (x86)
- ~1 EB of storage (50:50 disk/tape)
- Global networking (10-100 Gbps)

Ограничения иерархической модели MONARC



- Определение вычислительного ресурса, как соотношения вычислительных узлов, дискового пространства и систем архивирования информации, без учета WAN
- Статическая метод распределения данных между центрами;
- Методика обработки данных при статическом характере организации вычислительного ресурса и распределения данных («задания идут к данным»);
- Отсутствие понятия «популярности» (востребованности) для классов и наборов данных;
- Предположение гомогенности используемого ресурса и наличие ПО промежуточного уровня («middleware») во всех ВЦ.

Основное ограничение модели MONARC : иерархия центров и статический характер связки 1:T1-n:T2

- любой сбой в работе центра уровня Т1 практически останавливал работу всех связанных с ним центров уровня Т2;
- многие центры уровня Т2 были мощнее и стабильнее центров уровня Т1, но ресурс уровня Т2 не всегда мог быть использован оптимальным образом;
- ограничения при передачи данных. В рамках модели результаты выполнения заданий всегда д.б. быть переданы в центр уровня Т1 и только после этого могло быть создано дополнительное количество копий

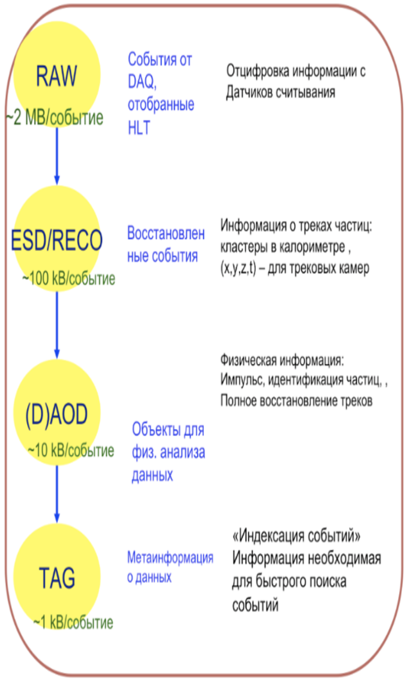
Разработка новой компьютерной модели. Переход от иерархической модели к «смешанной» модели в рамках грид инфраструктуры

При разработке новой модели были введены следующие определения :

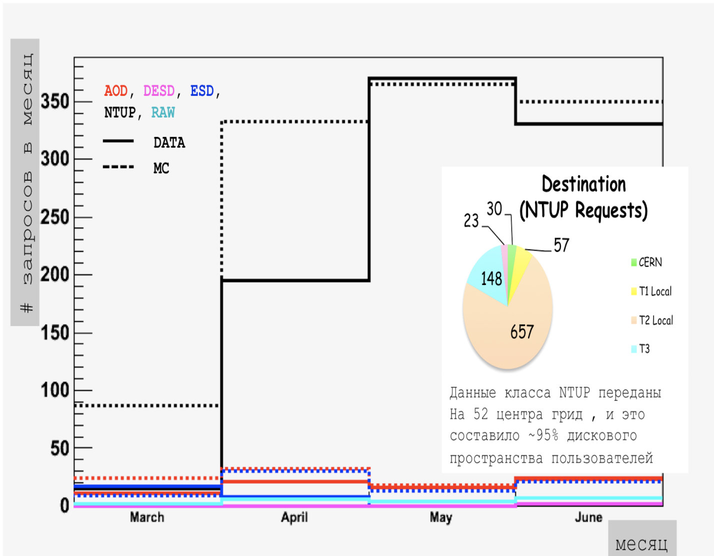
- **Популярность классов данных**. Насколько данные различных форматов популярны (востребованы) у ученых, и научных групп;
- **Температура данных**. Как со временем меняется частота обращения к наборам данных;
- **Вычислительная среда** : вычислительный ресурс, дисковый ресурс и ресурс архивирования, пропускная способность и стабильность глобальной вычислительной сети;

• В новой модели отсутствует предопределения функций центров;
  – Новый метод оценки стабильности работы центров, и как результат решение об использовании их дискового ресурса в качестве постоянного или временного, независимо от уровня центра в классификации WLCG.

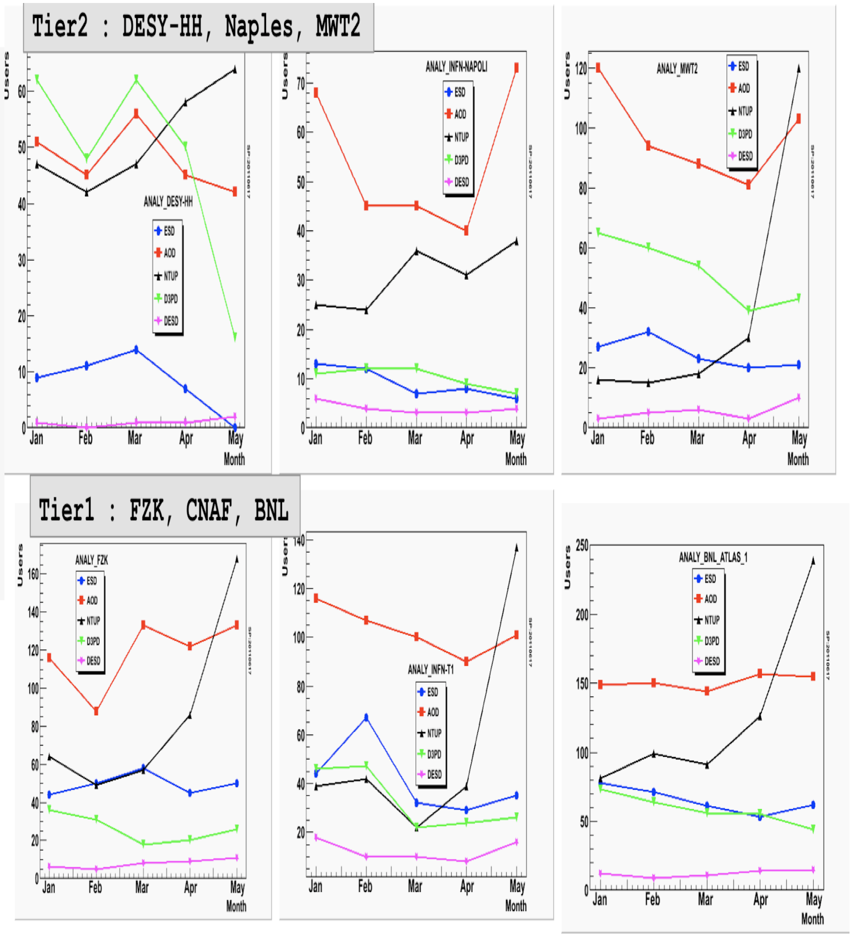
Классификация данных и их “популярность”



Шаги обработки данных в ФВЭ и ЯФ



Количество запросов пользователей на создание дополнительных копий наборов данных в распределенной инфраструктуре



Количество обращений задач анализа данных к различным классам данных в центрах уровня T1 и T2

Термодинамическая модель данных

Термодинамическая модель данных.

Была введена температурная шкала для всех данных, согласно показаниям шкалы состояние (Т) данных м.б. :

«Горячие данные». *Данные, широко используемые учеными и физическими группами эксперимента для проведения физического анализа и для исследования работы детектора.*

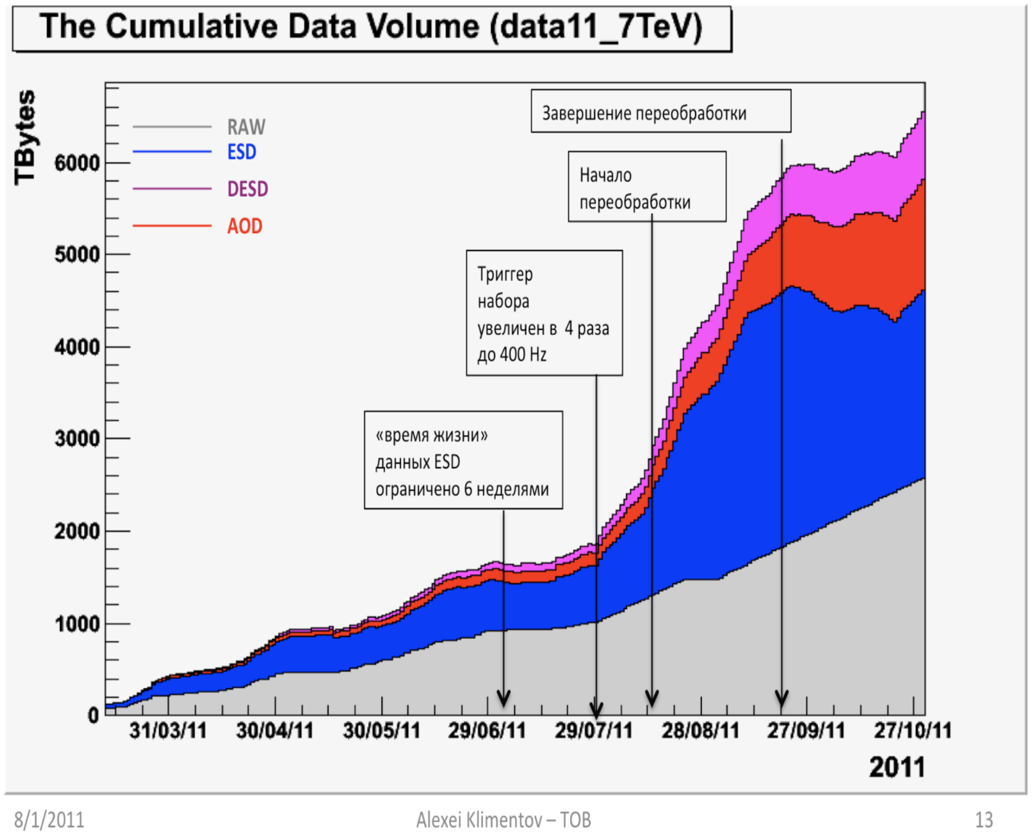
"Теплые данные". *Данные, используемые отдельными физическими группами и учеными для физического анализа. Предполагается, что уже имеется достаточное количество копий этих данных в грид инфраструктуре.*

"Холодные данные". *Данные, используемые отдельными физиками или рабочими группами. Две полные копии "холодных данных" распределены между центрами грид инфраструктуры;*

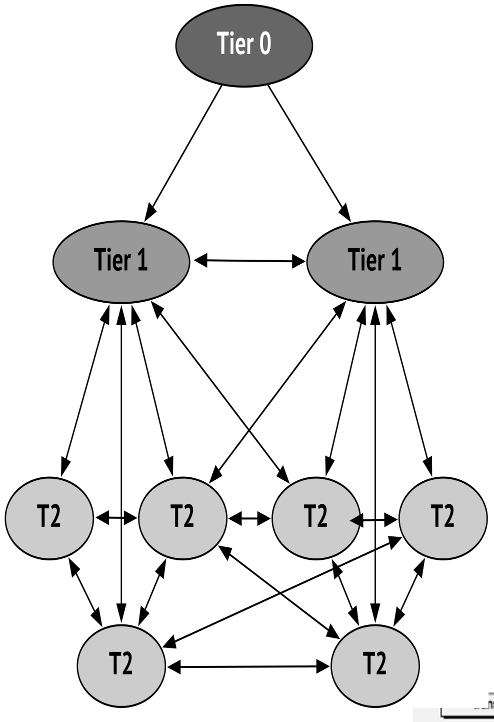
"Замороженные данные". *Данные «практически» не используются. Сохраняется одна полная копия (в центре, где данные были произведены или повторно обработаны).*

«Устаревшие данные» *Данные не могут быть использованы для физического анализа или других исследований в эксперименте, они подлежат удалению со всех сайтов и из всех каталогов.*

Данные эксперимента ATLAS в 2011 году. После этапа переобработки и введение термодинамической модели



Метод определения стабильности работы центров грид. Переход к «смешанной модели» компьютинга для экспериментов на БАК

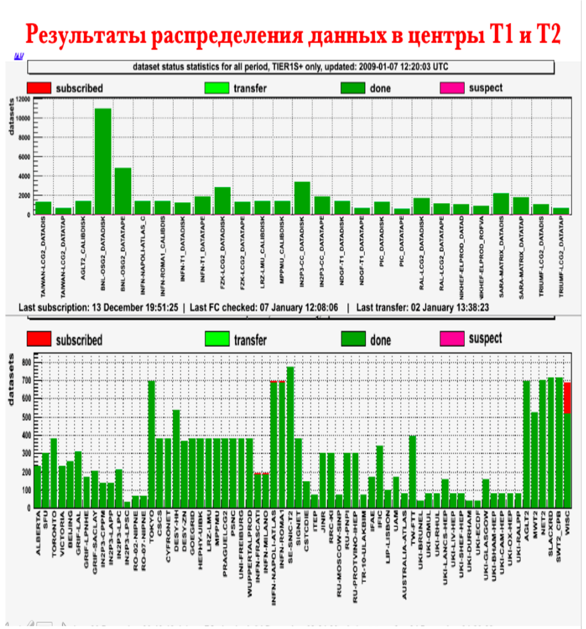


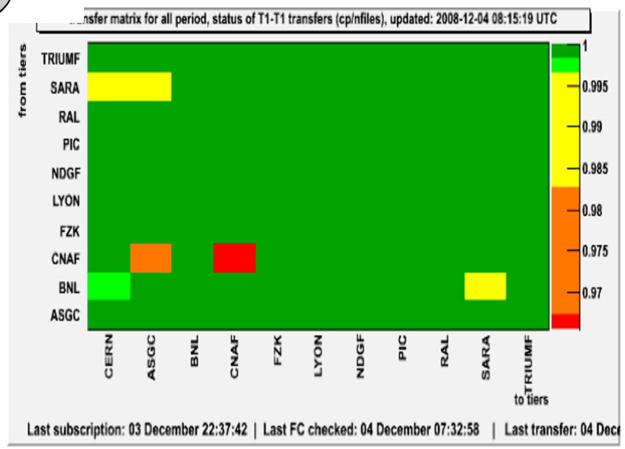
«смешанная модель» компьютинга для грид инфраструктуры

Были введены три независимые метрики :

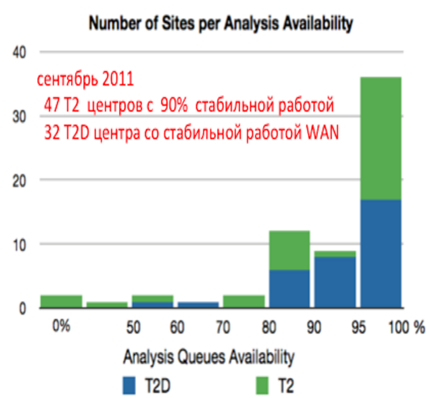
- Стабильность центра при обработке и анализе данных;
- Стабильность центра при обмене данными с другими центрами;
- Пропускная способность центра;

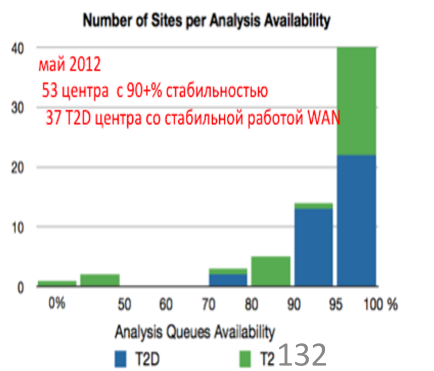
и четыре градации центров для их классификации : альфа, бета, чарли, дельта (от английского A, B, C, D). На первом этапе все центры уровня T2 были причислены к группе ‘альфа’.





Tier-2 Центры эксперимента ATLAS и их классификация





Метод динамического распределения данных с использованием информации о популярности данных

Статическое распределение данных для центров Т2 прекращается;

Дополнительные копии наборов данных создаются автоматически, следуя алгоритму :

Если задача пользователя обращается к набору данных и нет копии набора в центрах уровня Т2, то первая такая задача выполняется в центре первого уровня (Т1), где всегда есть копия данных, одновременно автоматически посылается запрос в систему управления данными для создания дополнительной копии набора данных на одном из центров уровня Т2;

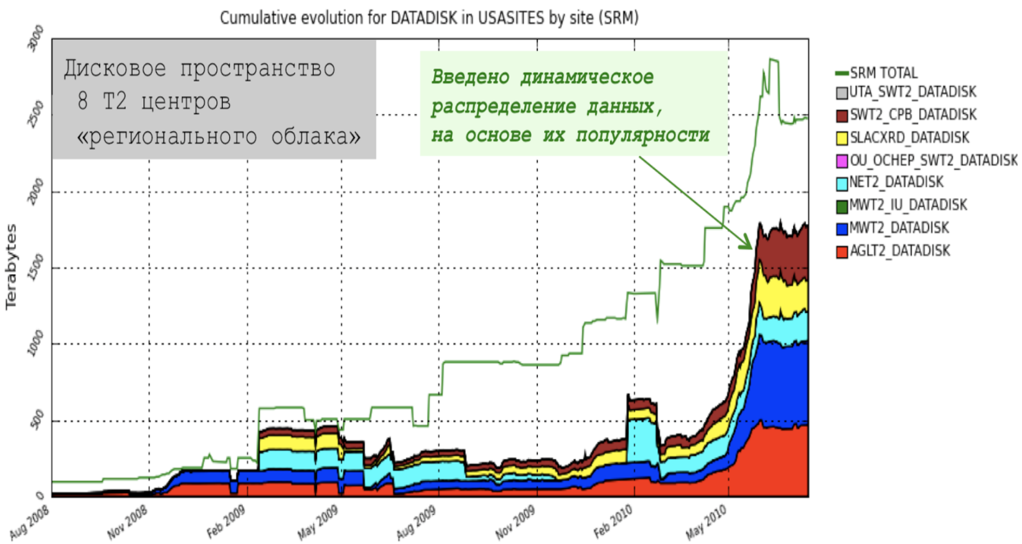
Таким образом для первого обращения к данным не существует задержки с выполнением задачи пользователя, а время передачи данных и создание дополнительной копии набора данных занимало несколько часов.

- Критерии выбора центра Т2 при создании дополнительной копии :
  - Свободное дисковое пространство;
  - Количество задач в очереди на выполнение к данному сайту;
  - Планируемая остановка сайта;
  - Количество файлов на передачу к данному сайту;

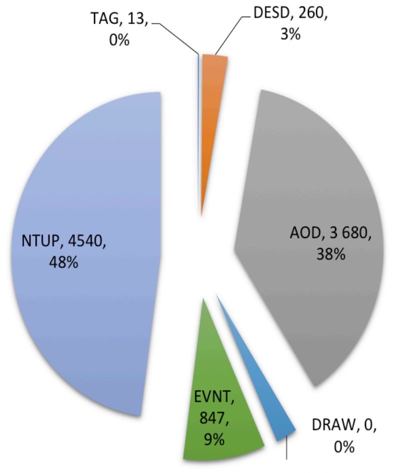
Общее количество копий данных основано на статистике обращения к ним задач пользователей и увеличивается логарифмически по мере роста количества обращений, т.е 10, 100 , 1000,... обращений соответствуют 1,2,3,... дополнительным копиям данных;

Данные форматов не используемых для физического анализа исключаются из рассмотрения;

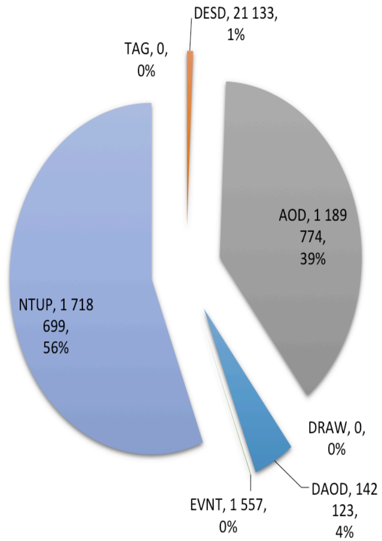
Метод динамического распределения данных с использованием информации о популярности данных

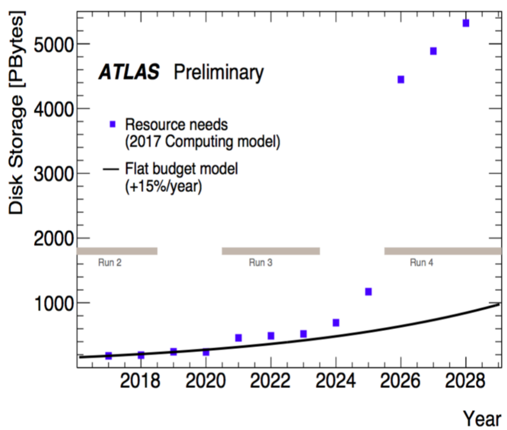


Количество наборов данных, для которых динамически было увеличено число копий



2+ использование дополнительной копии для разных классов данных





Потребности в дисковом пространстве
для второго и последующих этапов работы БАК