Title: Collider Experiments 2

Date: Jul 11, 2018 04:30 PM

URL: http://pirsa.org/18070031

Abstract:

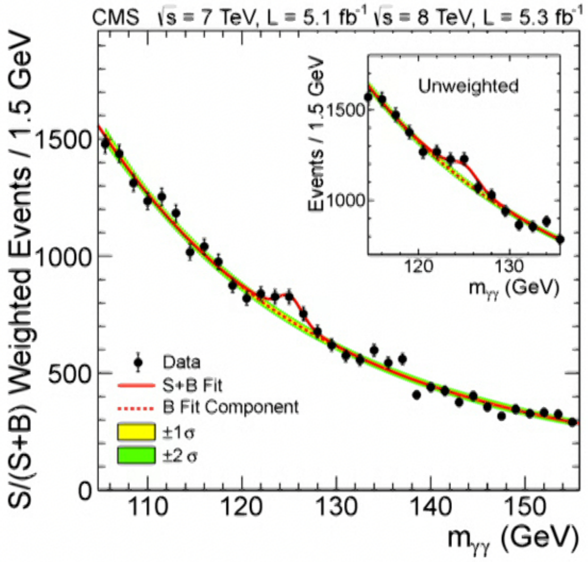
# Colliders Two:

Tom LeCompte

*High Energy Physics Division*
*Argonne National Laboratory*

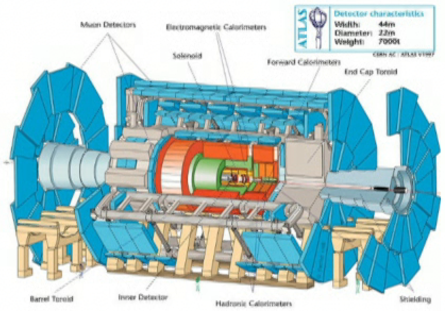
# A Question Over Lunch

- "Can you describe the role of Monte Carlo in the Higgs search?"
- The Higgs search in the discovery era was designed to use as little Monte Carlo as possible.
- Plots like this have the background fit from the data, not Monte Carlo.



# How It Works

Real Data Chain



Simulated Data Chain

**Event Generation**

Simulates the physics process of interest: produces lists of particles and their momenta

**Simulation**

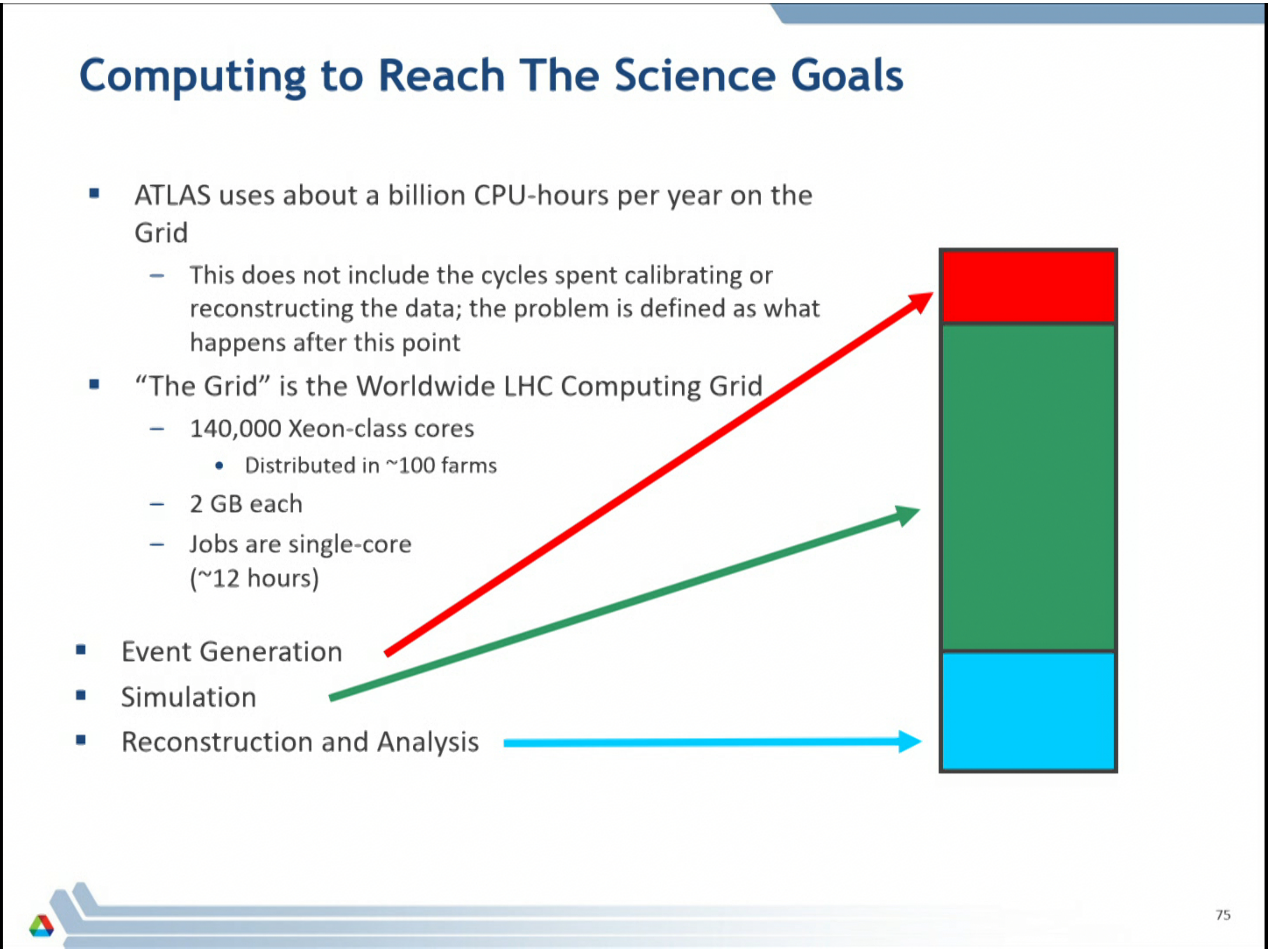
Simulates the interaction of these particles with the detector

**Reconstruction**

Infers particles that must have been present based on the detector response

**Reconstruction**

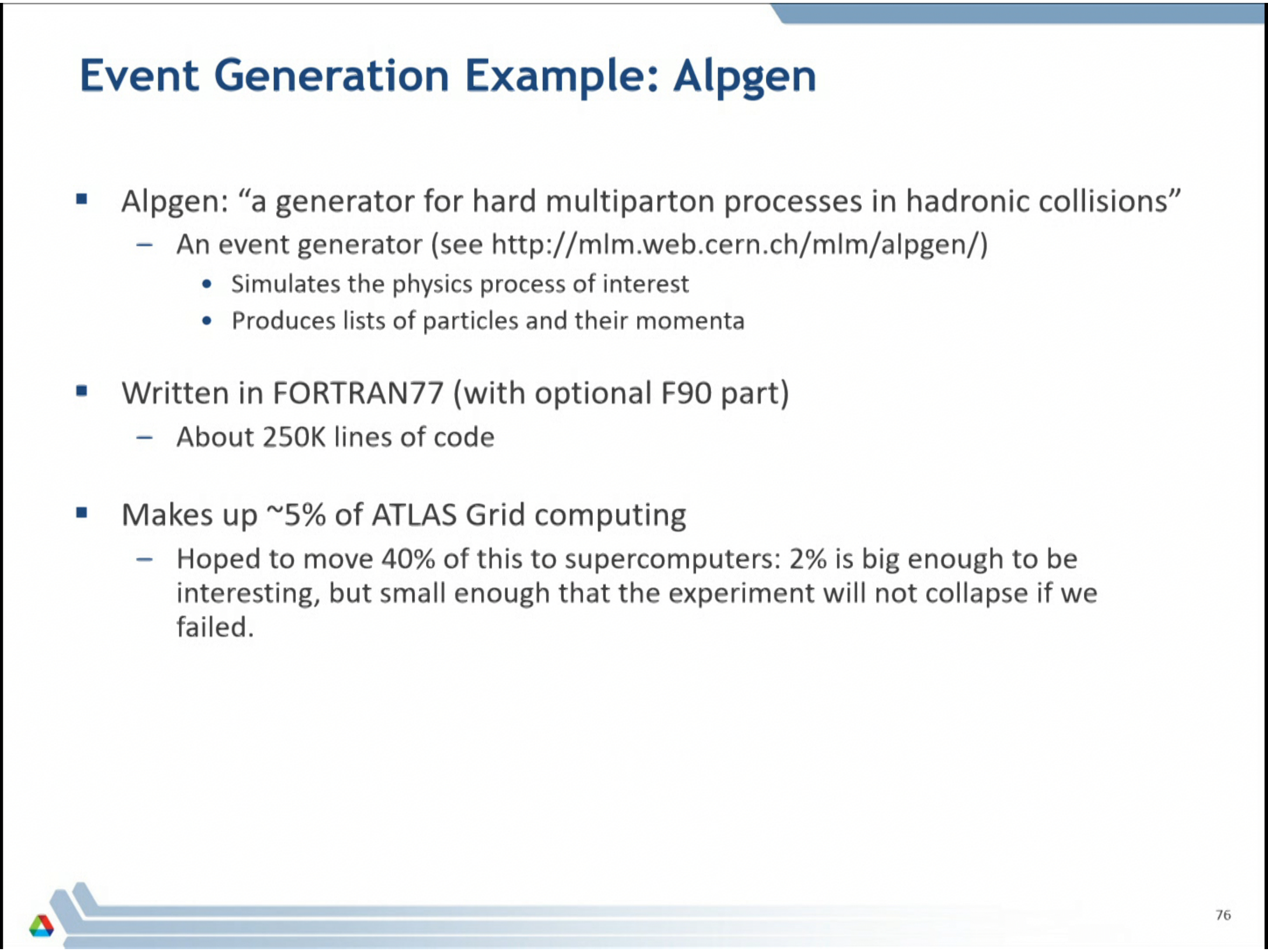
Analysis: comparing the two



# Computing to Reach The Science Goals

- ATLAS uses about a billion CPU-hours per year on the Grid
  - This does not include the cycles spent calibrating or reconstructing the data; the problem is defined as what happens after this point
- "The Grid" is the Worldwide LHC Computing Grid
  - 140,000 Xeon-class cores
    - Distributed in ~100 farms
  - 2 GB each
  - Jobs are single-core (~12 hours)

- Event Generation
- Simulation
- Reconstruction and Analysis

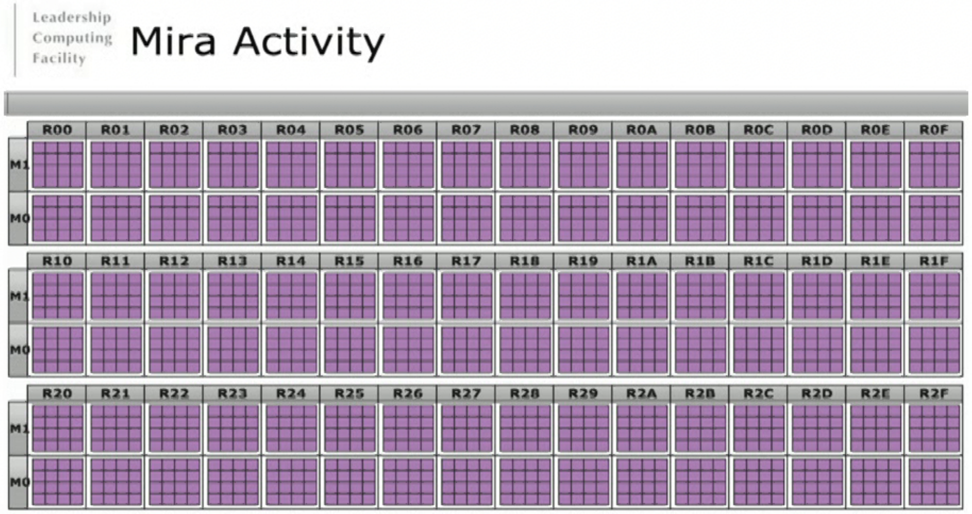


# Event Generation Example: Alpgen

- Alpgen: “a generator for hard multiparton processes in hadronic collisions”
  - An event generator (see http://mlm.web.cern.ch/mlm/alpgen/)
    - Simulates the physics process of interest
    - Produces lists of particles and their momenta
- Written in FORTRAN77 (with optional F90 part)
  - About 250K lines of code
- Makes up ~5% of ATLAS Grid computing
  - Hoped to move 40% of this to supercomputers: 2% is big enough to be interesting, but small enough that the experiment will not collapse if we failed.

# Supercomputing and Monte Carlo

- Mira is a supercomputer at ANL (#17 in the world in 2018)
  - ¾ of a million processors
  - Combined computational power of all the cell phones in Toronto
- Based on this success, the experiment asked us to do all the Alpgen gheneration for the next two years

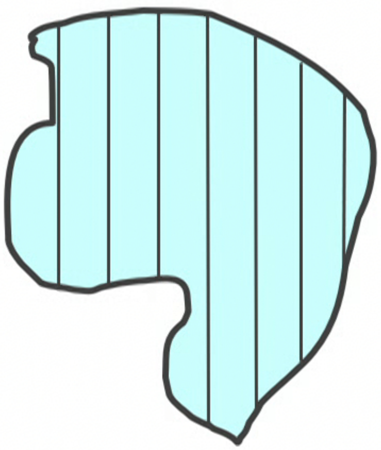

Leadership Computing Facility

Mira Activity

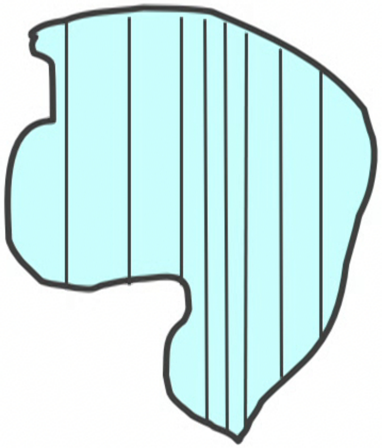
While this job was running, Mira was producing the equivalent computing as 5 or 6 ATLAS Grids.

On our best days, we provide the equivalent computing capacity of the whole ATLAS Grid.

# Phase-space integrals



You can have each rank work on part of the problem, and sum them later.  Equal spacing means some ranks finish sooner than others.

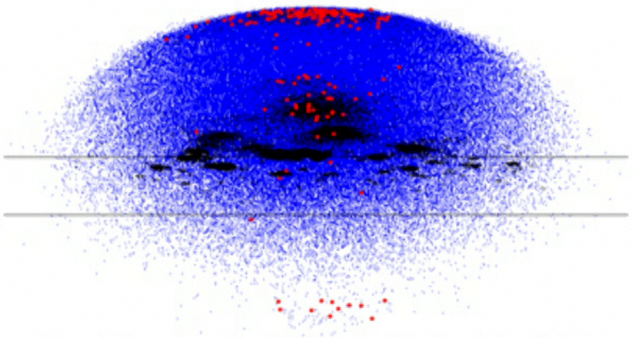


You can better balance things with unequal spacing, trying to match areas.  There is a limit to this – if you knew the areas exactly, you wouldn't need to do the integral.

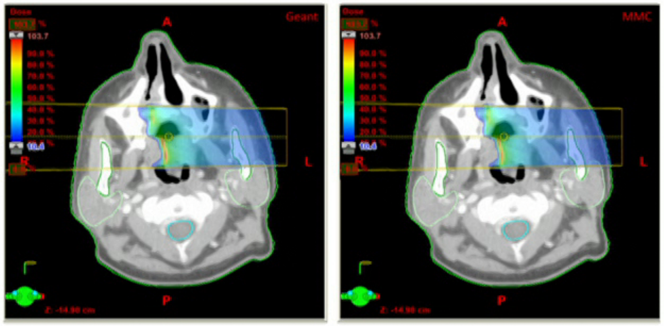
# Simulation/Digitization

- Simulates the interaction of these particles with the detector
- This is the green part of the diagram
  - Once ~70% of the usage
  - Falling to 60% or even less – but only because the red (generation) has grown
- This uses a C++ toolkit called Geant4
  - http://geant4.cern.ch/
  - About 1 million lines of code
  - The only program HEP uses for this
    - Modular, so different models can be inserted in the same framework
  - Identifies the position and amount of energy deposits in the detector
    - Energy is all that matters

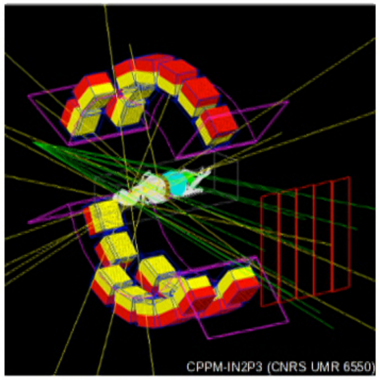
# Geant4 goes beyond HEP


Electron avalanches in thunderstorms


Radiation effects on satellites

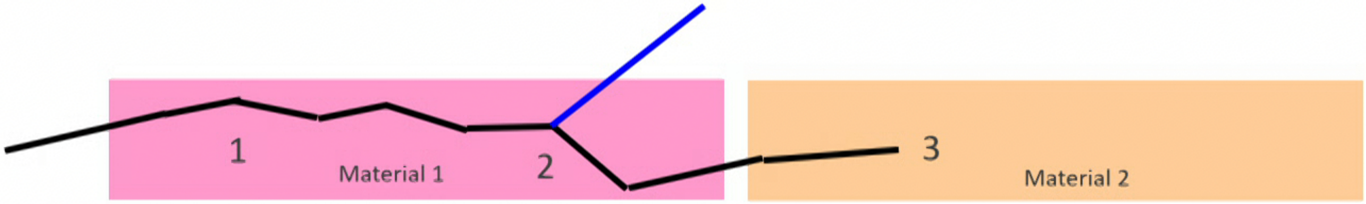


Dose calculations in proton therapy


CPPM-IN2P3 (CNRS UMR 6550)

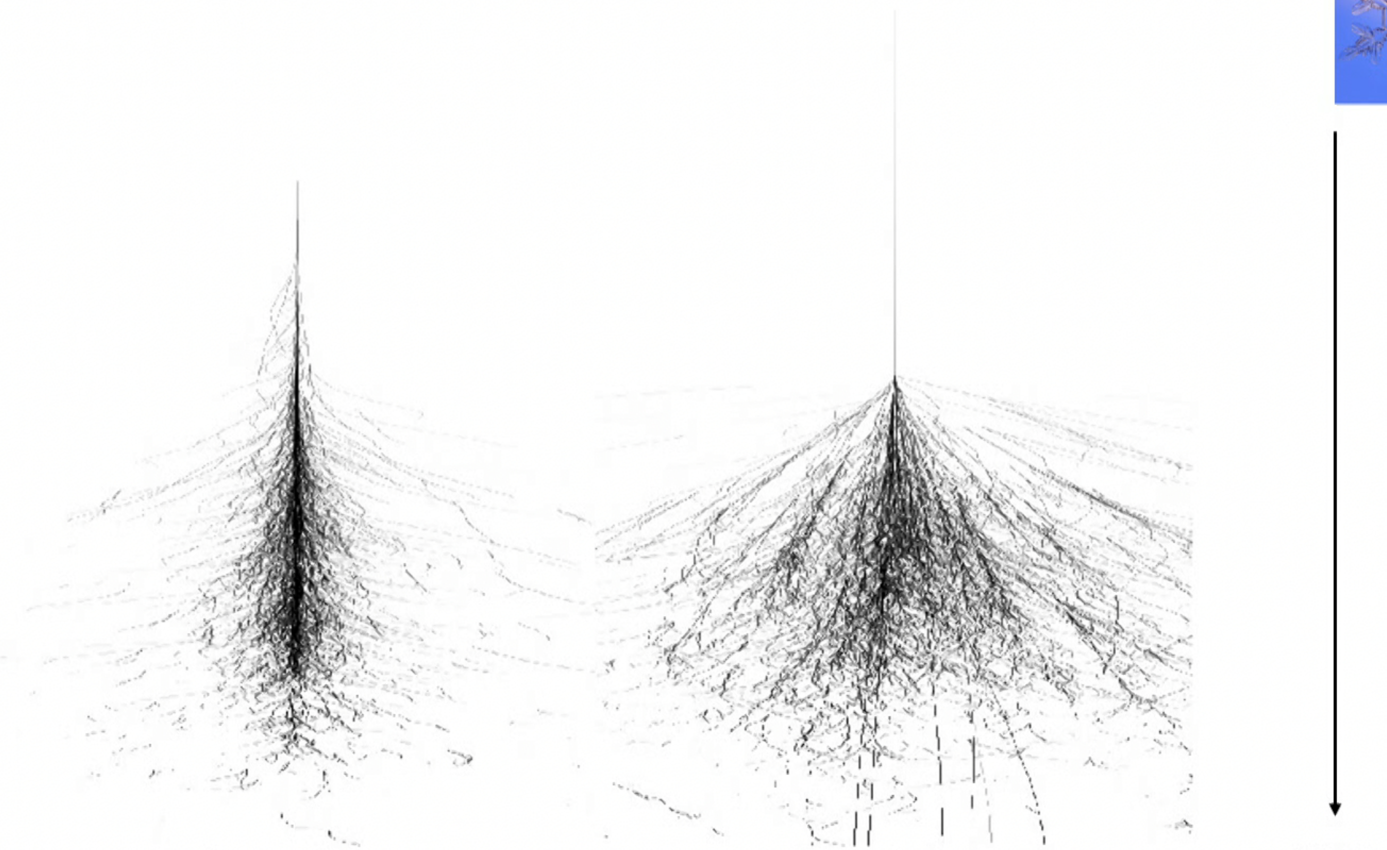
Improved PET scanning development

# How It Works



- The program takes small steps with each particle through the detector. At each step, one of three things can happen:
  - 1. The particle deposits some energy via ionization, and takes another step
  - 2. A new particle is created, and added to the list of particles
  - 3. The particle is absorbed, stopped, or exits the detector and removed from the list
- Geant starts with the particles produced in the collision and continues until the particle list is empty
- This takes a few minutes per event (for a typical ATLAS event)

# Hadronic Showers



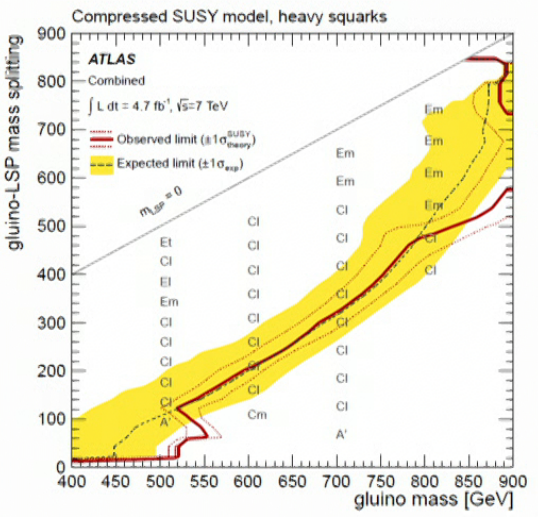
Gamma shower

Hadronic shower

Muon

# Aside: How I Got Into Supercomputing

- When I came back after being ATLAS physics coordinator, I was looking for some physics to do.
- Steve Martin (NIU) and I asked ourselves about the sensitivity of the LHC experiments to moderately compressed SUSY spectra
  - Answer: better than you might think (c.f. PRD84 015004 and PRD85 035023)
- ATLAS did the study, and here's the outcome (as shown in the 7 TeV paper):



The problem

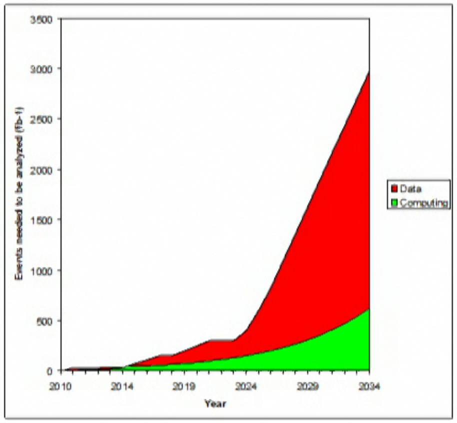
We didn't simulate these points, so we have no idea what the limits are here.

- The observed limit is better than our expected limit. We (perhaps foolishly) did not run any points out there.
- This set of points spent ~2 months in the queue waiting for open slots.
- We decided not to delay this paper by another ~2 months

# High Performance Computing Motivation in HEP

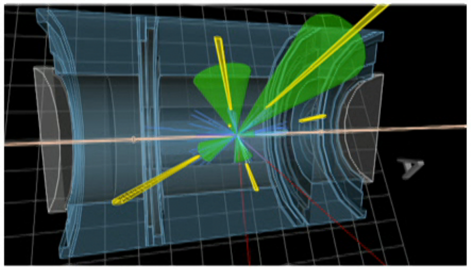
There are two problems we are trying to solve

- The capacity problem
  - Our needs are growing faster than the Grid is growing – and likely can grow



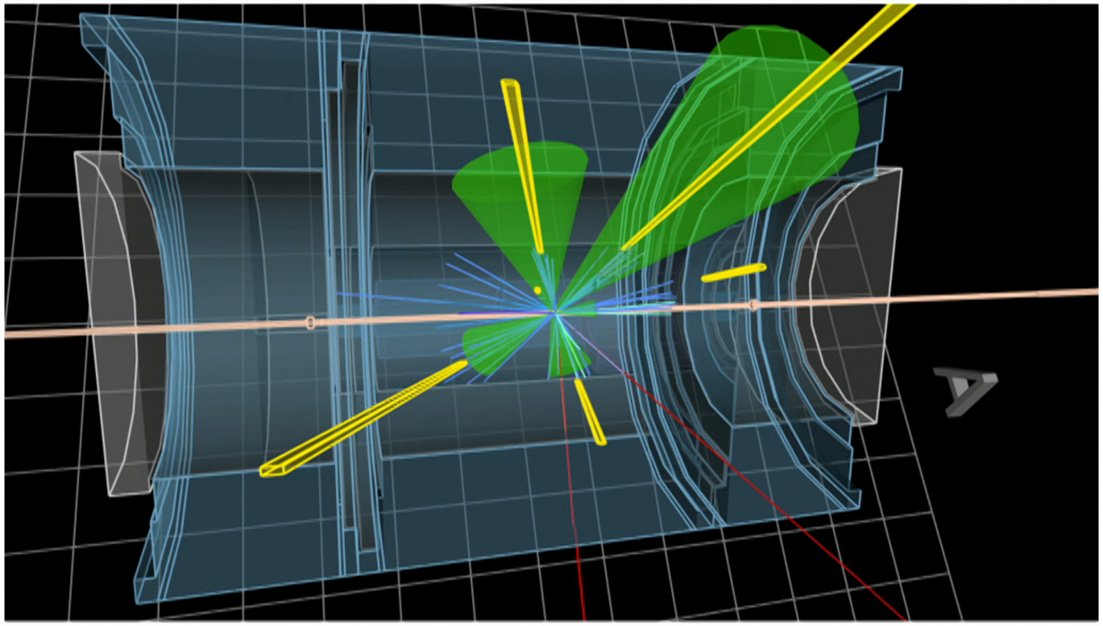
The green assumes 15% growth per year from Run 1, and that Run 1 had exactly enough capacity.

- The capability problem
  - Some problems can't run on the Grid
    - Highly filtered samples can produce zero events in a job – which looks like a failure and triggers a resubmit. Which triggers a resubmit. And so on.
    - Sherpa integration times stretch into the weeks, but can be done overnight with a few hundred threads (scaling is poor today, but this makes doing this *possible*. Efficient comes later)



We see HPCs playing a role in both of these

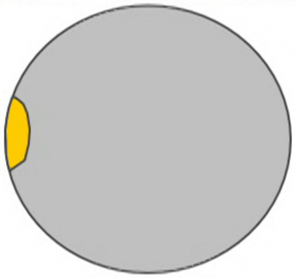
# Events Like These:



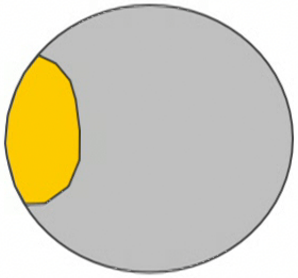
- This is a Z (→ ττ) + 5 jet event, with highly filtered τ decays.
- ATLAS requested that we make them, because they failed on the Grid
  - The degree of filtering is so high, runs often have zero events pass – misinterpreted as a failure.   This is not a problem with thousands of cores.
- There are another 46,998 events just like this one.

# What NOT to Do With Your Magnet

Suppose a small region in your superconducting cable goes normal.



Current will flow around the resistive spot, driving it past $j_c$; the spot grows



Eventually, the entire cross-section goes normal, and now you have a resistive wire. All the heat is dissipated in that spot.

Stored energy in magnets = 10 GJ, same as a 747 at top speed.)

A magnet undergoing a controlled quench.

# 2008 “Incident” Timeline

- March 2008 – CERN announces the LHC will start with 5 TeV per beam rather than 7 TeV. This avoids a lengthy magnet retaining process.
- 10 September 2008 – amidst much media hoopla, beams are circulated at 450 GeV (injection energy). At this time, 7 of the 8 sectors are “qualified for 10 TeV collisions”, meaning they operate properly at 11 TeV equivalent current.
- 18 September 2008 – a transformer near Point 5 fails. EDF says it will take a couple of days to find and install a replacement. Two sectors start to warm. Decided to return to qualifying the last sector, 3-4, in parallel.
- 19 September 2008 – during one of these tests, a magnet quench led to an electrical arc, which in turn led to a catastrophic loss of helium, which made a great big mess.

# Aftermath

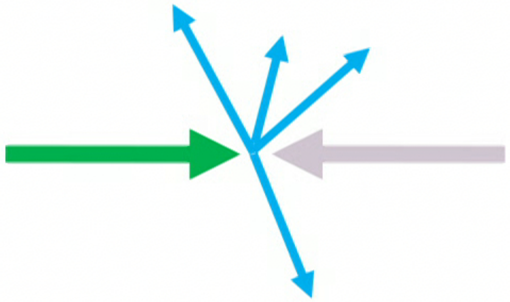
- A one-year delay
  - 53 magnets had to be removed and repaired or replaced
  - Quench protection system redesigned and replaced
  - Additional vents were added to most of the dipoles
- Beam energy lowered to 3.5 TeV per beam in 2011
  - This is 4x safer ($P=I^2R$)
  - This also allows for a lower dump resistance and faster dump – reduced risk.
  - However, the physics reach is less (by 5-10), but still far beyond what we had before.
  - Energies 4.0 TeV in 2012, 6.5 TeV in 2015

# Why Build A Collider Anyway?

Before colliders, we had fixed-target experiments

Beam hit a target, reaction products go forward where they are studied.

- The more energetic the beam, the more momentum it carries
- The more momentum it carries, the more the reaction products go forward
- To look for new particles and new phenomena, you want the center of mass energy (a Lorentz-invariant quantity) as high as you can get it
- Increasing the beam energy causes the center of mass energy to grow only as the square root



- If instead we collide the beams, we can work in the center of mass frame
- The price we pay? Luminosity.

# Luminosity

$$L = N_b f \frac{N_1 N_2}{4\pi\sigma_x\sigma_y} = N_b f \frac{N_1 N_2}{4\pi\varepsilon\beta}$$

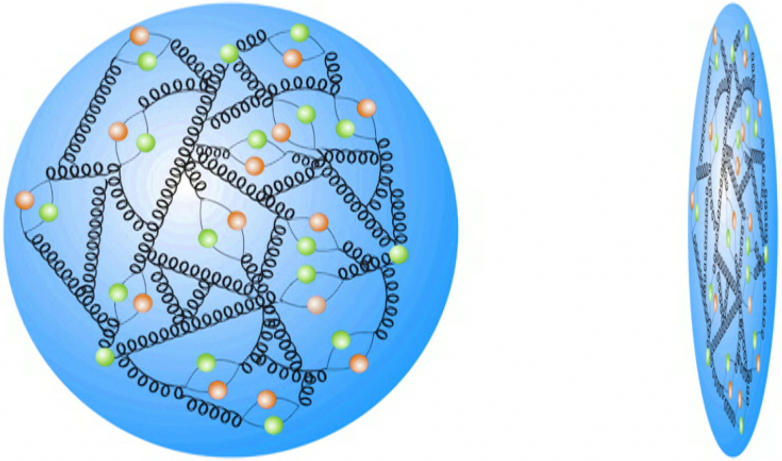
- Elements of the Equation
  - $N_b$ is the number of bunches in the ring } These are, of course, correlated
  - f is the frequency of revolution
  - $N_1$ and $N_2$ are the number of particles in each bunch
  - The denominator is the area of the beam
    - I can express it in lengths ($\sigma$)
    - But it's more useful to express it in terms of emittance ($\varepsilon$), a property of the beam itself and $\beta$, a property of the accelerator FODO lattice
      - For simplicity, I made the beam circular in the 2nd equation
    - We call $\beta$ at the interaction region $\beta^*$.
- Luminosity has the rather odd-looking units of inverse area per unit time
  - Why will become clear soon

# Getting Higher Luminosity

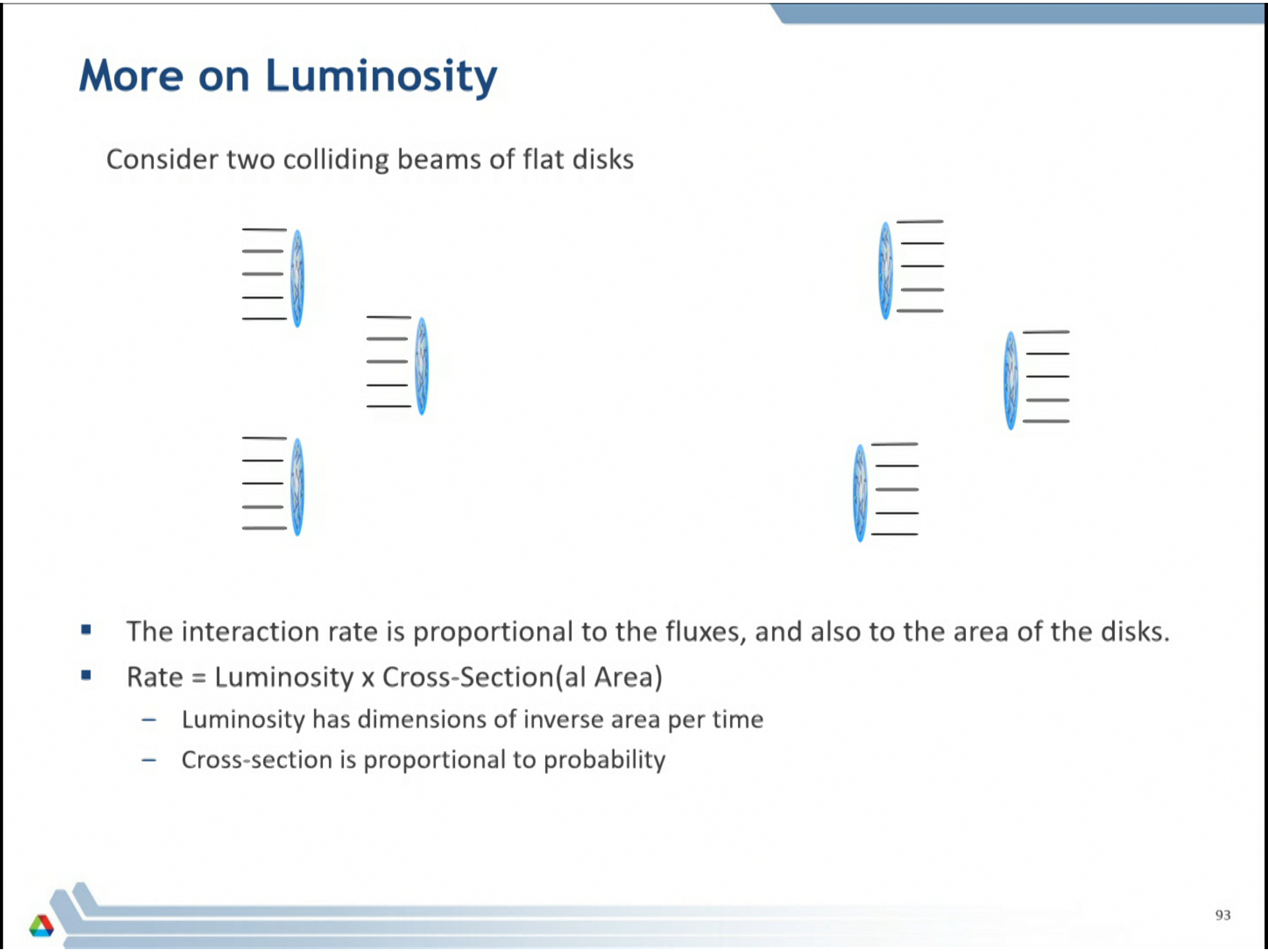
$$L = N_b f \frac{N_1 N_2}{4\pi\varepsilon\beta}$$

- What elements do we have control over?
  - $N_b$ is the number of bunches in the ring – fixed once you know the RF
    - A bit of a lie, because not every bucket is filled, e.g. the Tevatron had 1113 buckets but only 36 filled. For the LHC those numbers are 35640 and 2808 (at design).
  - f is the frequency of revolution – again fixed once you know the RF
  - $N_1$ and $N_2$ are limited by electrostatic repulsion of the beam
    - For the Tevatron $N_2$ was determined by the antiproton production rate
  - The emittance is a function of the beam quality in the injector chain
    - Small is good, but…
    - More beam usually means higher emittance
  - Lowering $\beta^*$ is good, but remember quadupoles are lenses
    - The smaller I make the beam at the interaction point,
    - the larger I make it everywhere else. There are limits.

# The Proton



- A proton is a very dynamic thing composed of interacting antiquarks, quarks (3 more than the antiquarks, and gluons.
- At LHC energies, $\gamma$ = almost 7000
  - This Lorentz contracts the proton into a thin disk
    - Aspect ratio about the same as a sheet of paper
  - Time dilation “freezes” the proton so it appears static
- A 4-dimensional object becomes a 2-dimensional object



# More on Luminosity

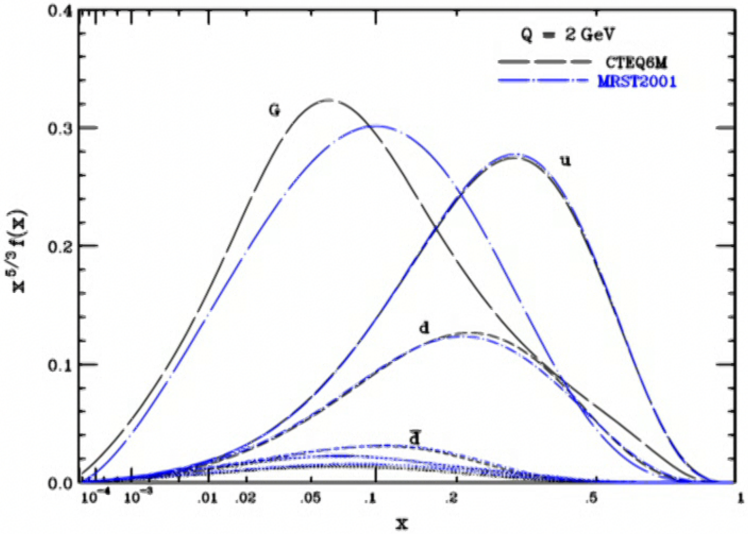
Consider two colliding beams of flat disks

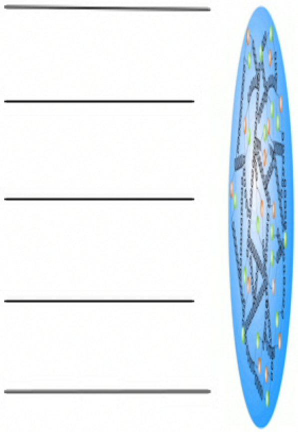
- The interaction rate is proportional to the fluxes, and also to the area of the disks.
- Rate = Luminosity x Cross-Section(al Area)
  - Luminosity has dimensions of inverse area per time
  - Cross-section is proportional to probability

# Cross-section, Luminosity and Units (by example)

- A square meter is too big to be practical
- Instead we use a conventional unit called a “barn” = $10^{-28}$ $m^2$
  - From the expression “broad side of a barn”
  - This is still too big to be practical, so we usually talk about millibarns, microbarns, nanobarns, picobarns, etc.
- Example: the top quark cross section at 7 TeV is 175 pb.
  - Easy to remember, since the mass is 175 GeV.
  - Since the proton total inelastic cross section at 7 TeV is 100 mb…
  - The probability of producing a top quark is (175 pb)/(100 mb) = $1.75 \times 10^{-9}$
- A luminosity of $10^{33}$ $cm^{-2}$ $s^{-1}$ works out to 1 $nb^{-1}/s$
  - 1 $nb^{-1}$ second x 100 mb = interaction rate of 100 MHz
  - 1 $nb^{-1}$ second x 175 pb = top quark rate of 6/minute

# The Frozen Proton



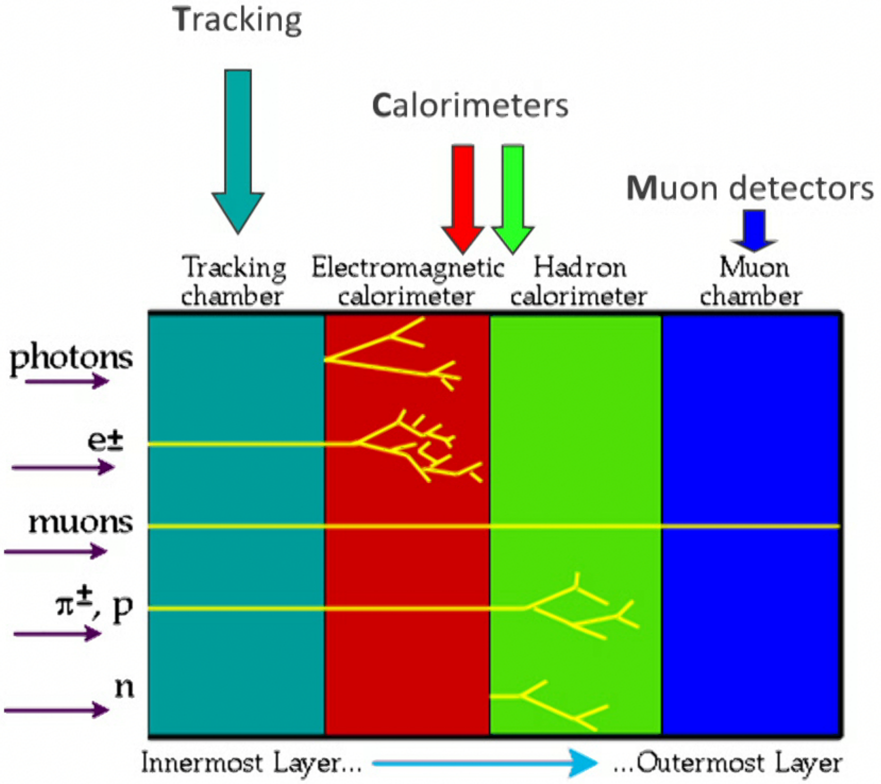


- Because the proton is frozen in time, we can envision each parton as carrying a fraction x (sometimes called Bjorken x) of the proton's total momentum

- We talk about the proton-(anti) proton center of mass energy, but what's really relevant is the parton-parton center of mass energy, which is
  - Variable
  - Smaller (around 20% the energy)

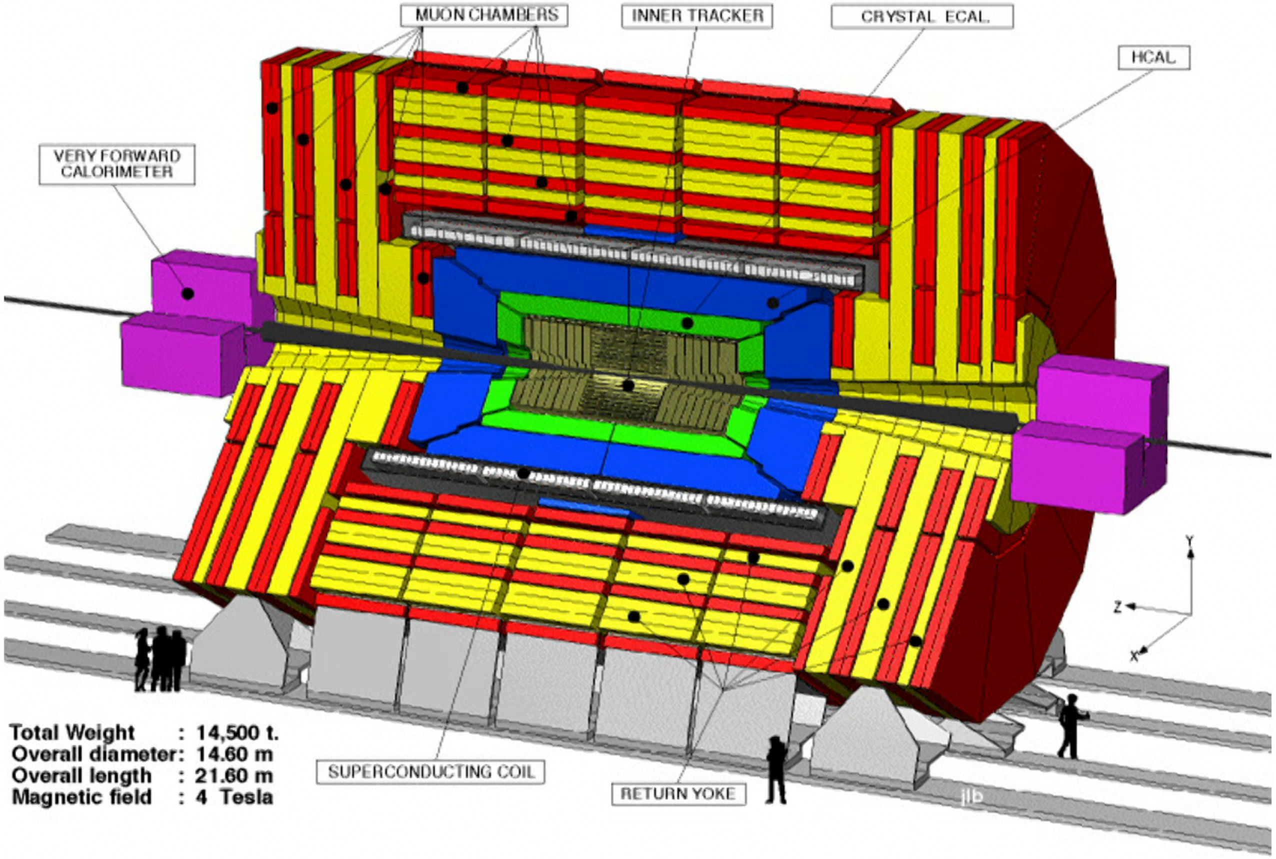
# Onto Experiment Design…

# Fact One: The basic design of experiments is the same:



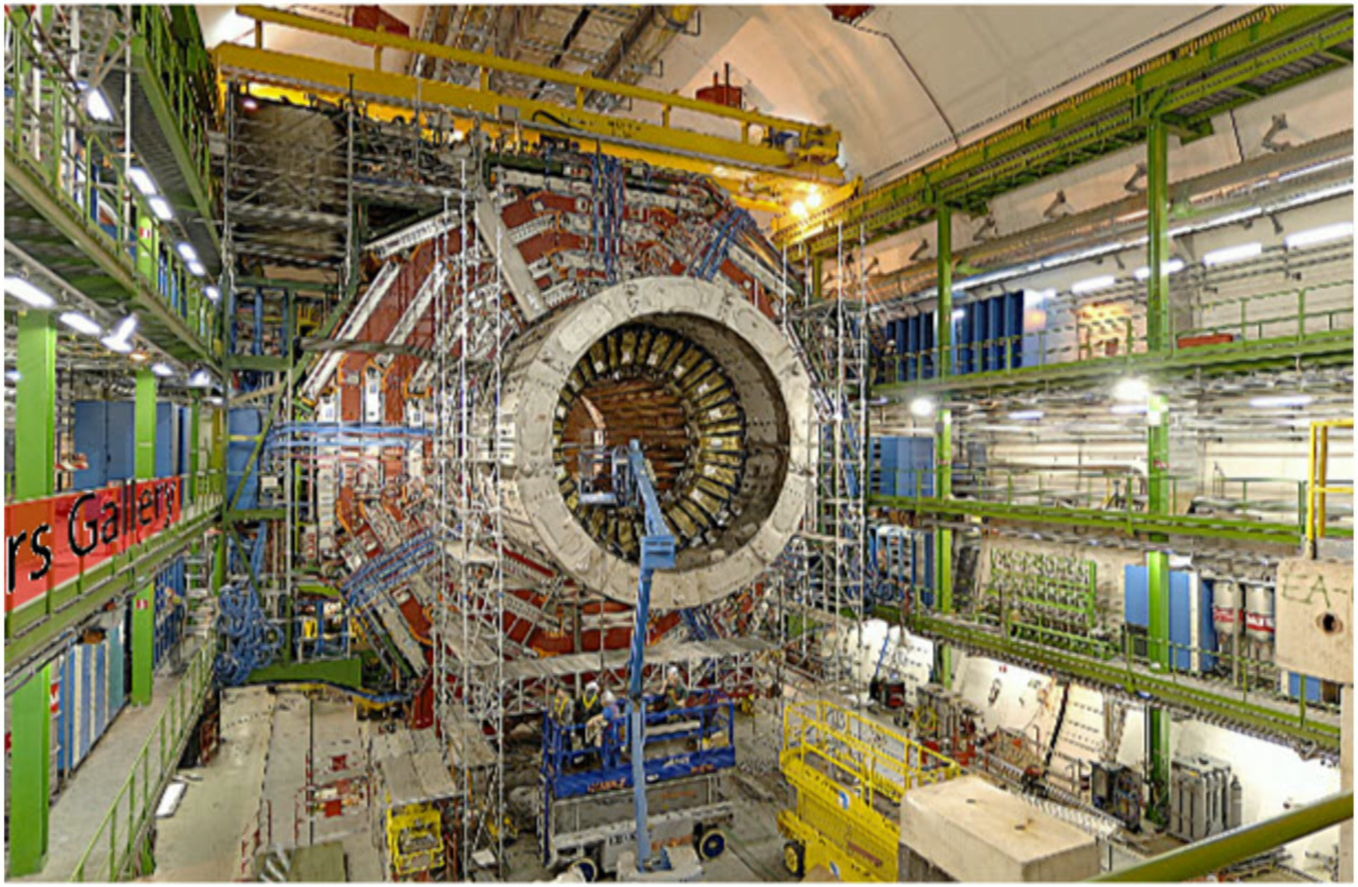
- Driven by the physics of the interaction of high energy particles with matter.
- Because the physics is the same, successful experimental designs are similar

# The Compact Muon Solenoid

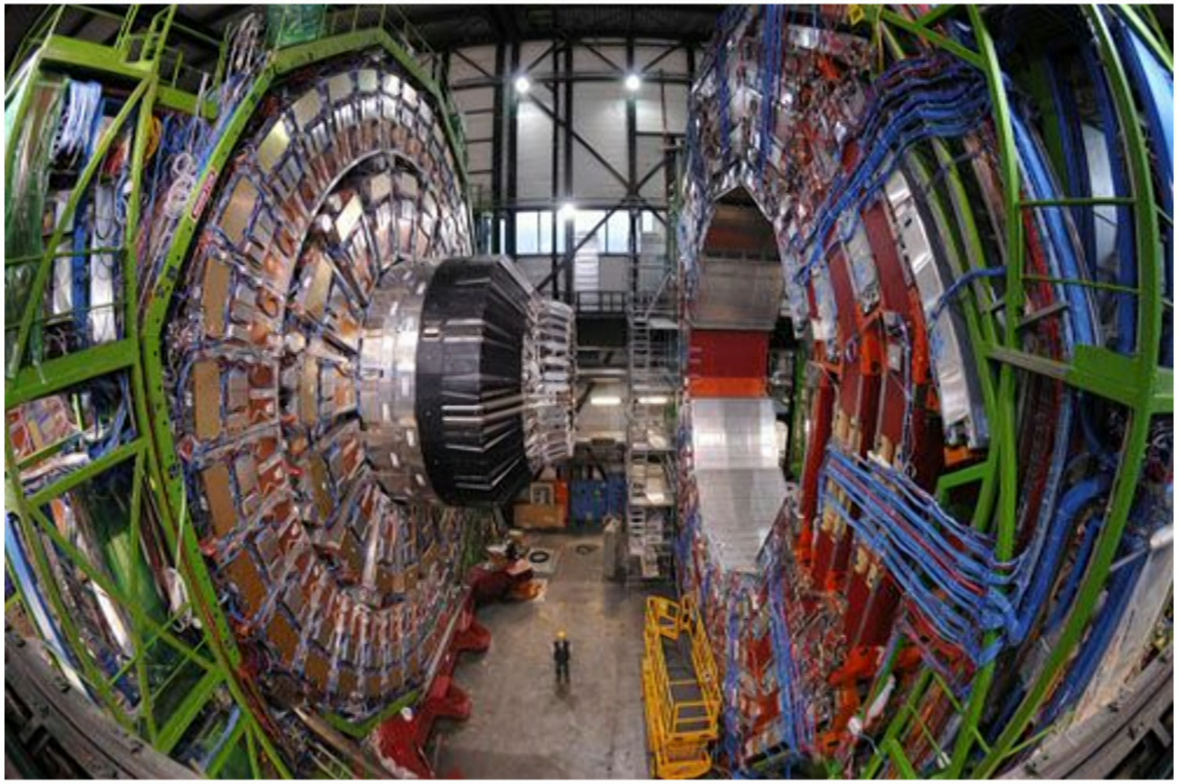


**Total Weight** : **14,500 t.**
**Overall diameter**: **14.60 m**
**Overall length** : **21.60 m**
**Magnetic field** : **4 Tesla**

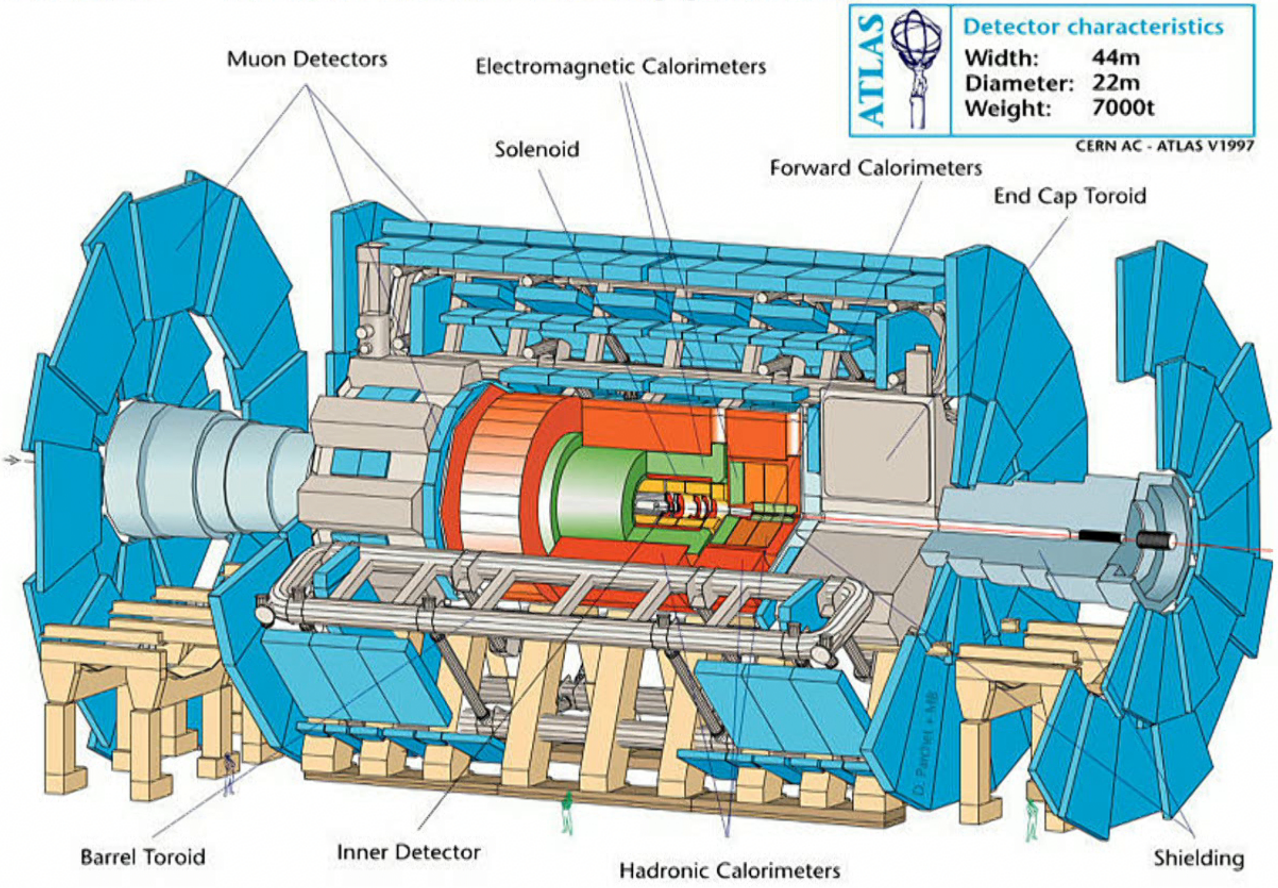
# CMS Under Construction



# CMS Today



# ATLAS = A Toroidal LHC ApparatuS

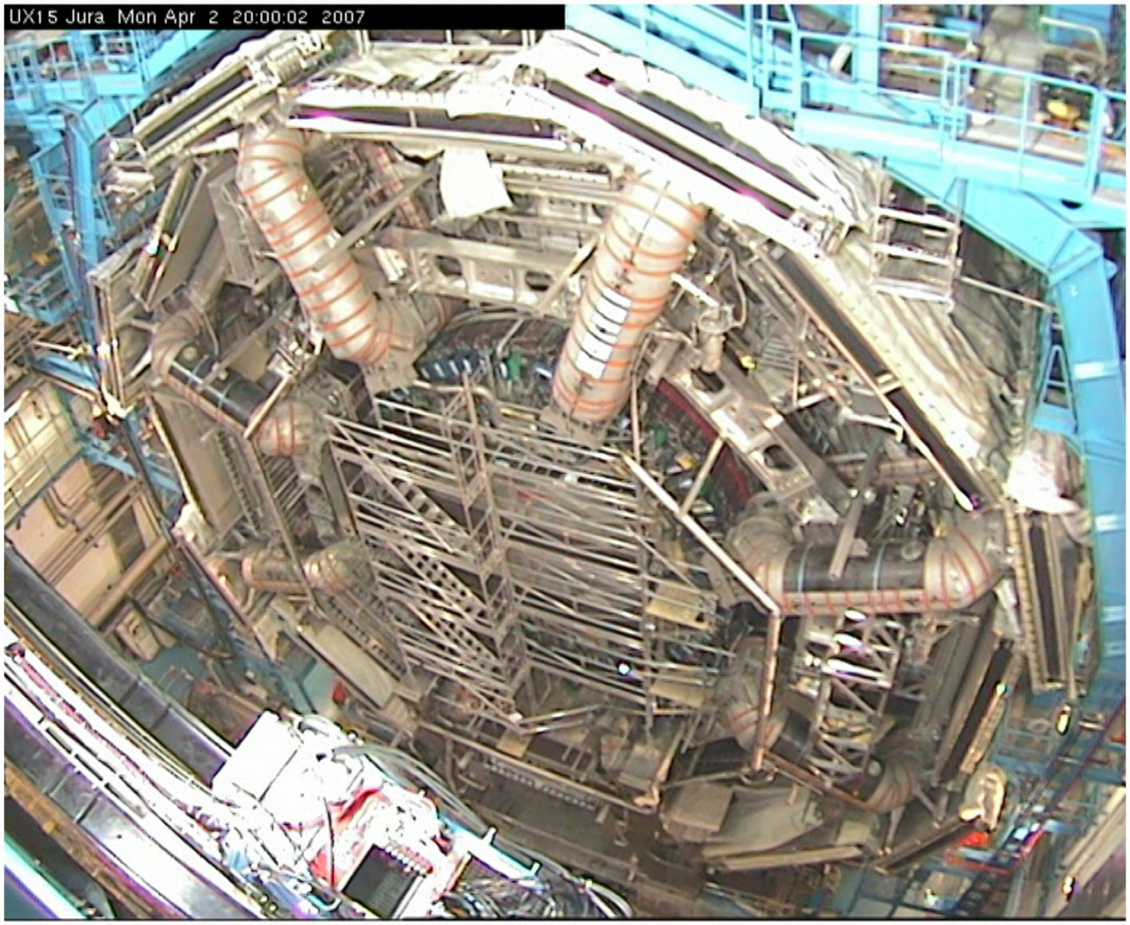


**Detector characteristics**

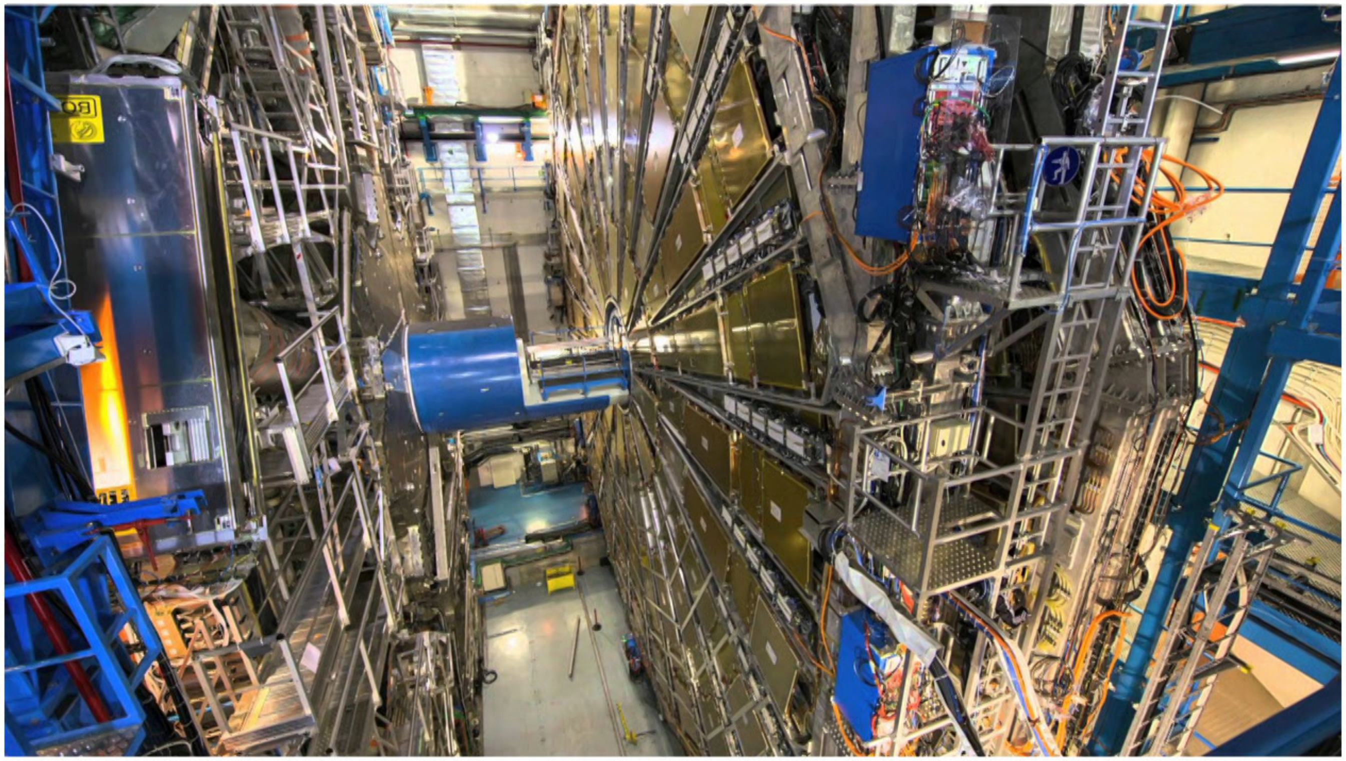
Width: 44m
Diameter: 22m
Weight: 7000t

CERN AC - ATLAS V1997

# ATLAS Under Construction

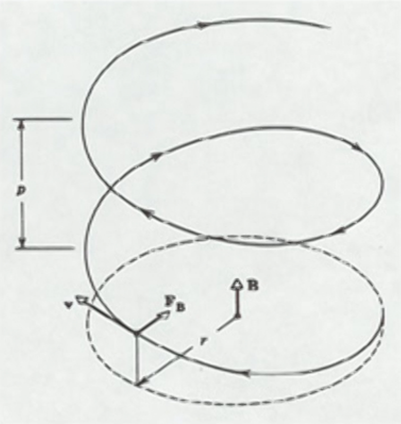


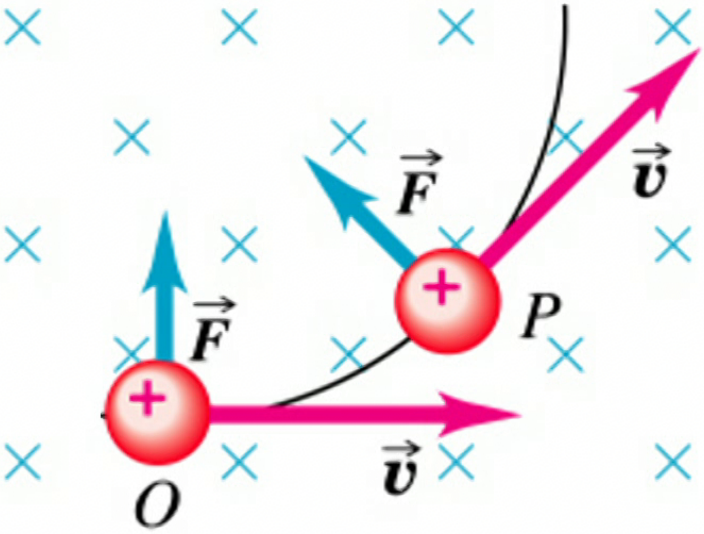
# ATLAS Today



# Fact Two: Tracking measures 1/p

Charged particles in a uniform magnetic field $\vec{B} = B_0\hat{z}$ move in helices:



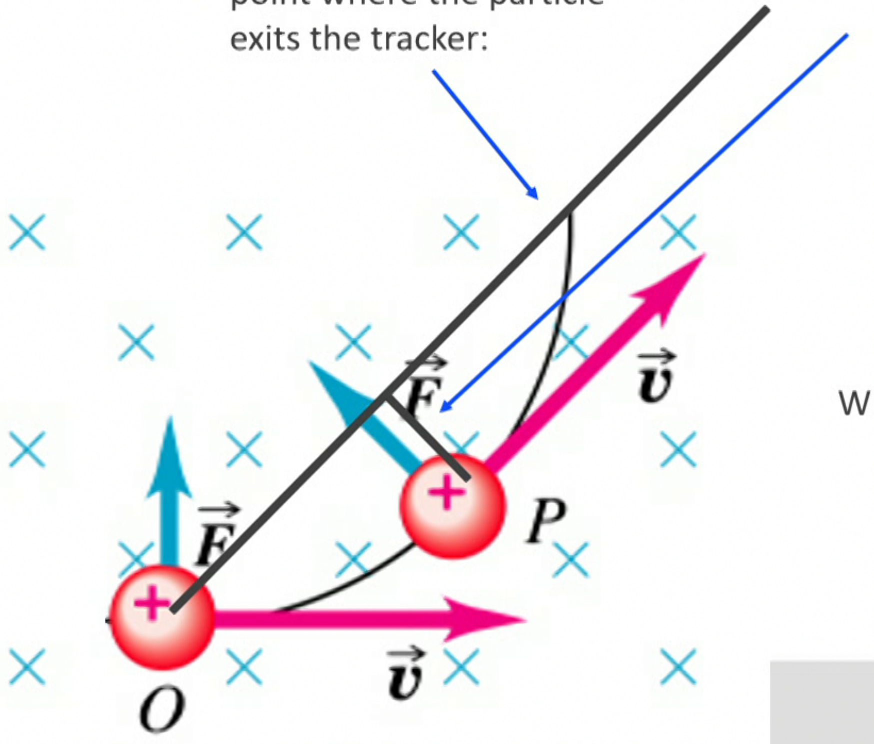


It's convenient to work in the transverse plane (i.e. the plane normal to the Z direction)

In this plane, the helices project to circles.

# Fact Two: Tracking measures 1/p (II)

Radial line from origin to point where the particle exits the tracker:



The sagitta ("arrow") $s$ is the distance of maximum deflection from a straight line track:

$$s = \frac{qBL^2}{8p_T}$$

or

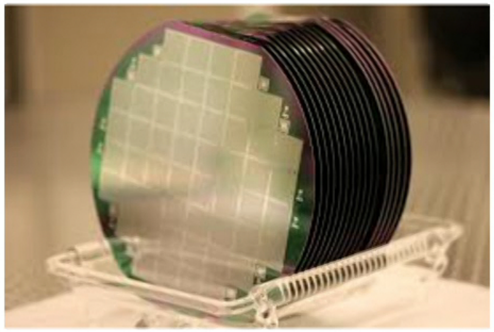
$$p_T = \frac{qBL^2}{8s}$$

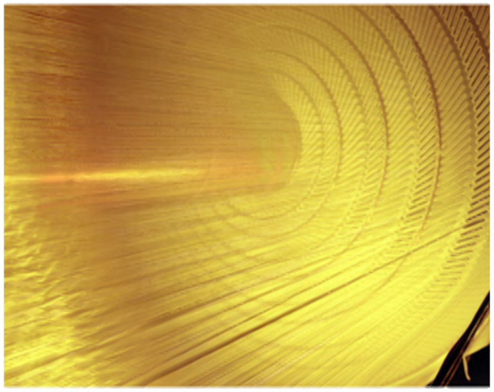
Which leads to the expression

$$\frac{\delta p}{p} = \frac{\delta s}{s} \propto p$$

As momentum increases, tracking becomes more difficult.

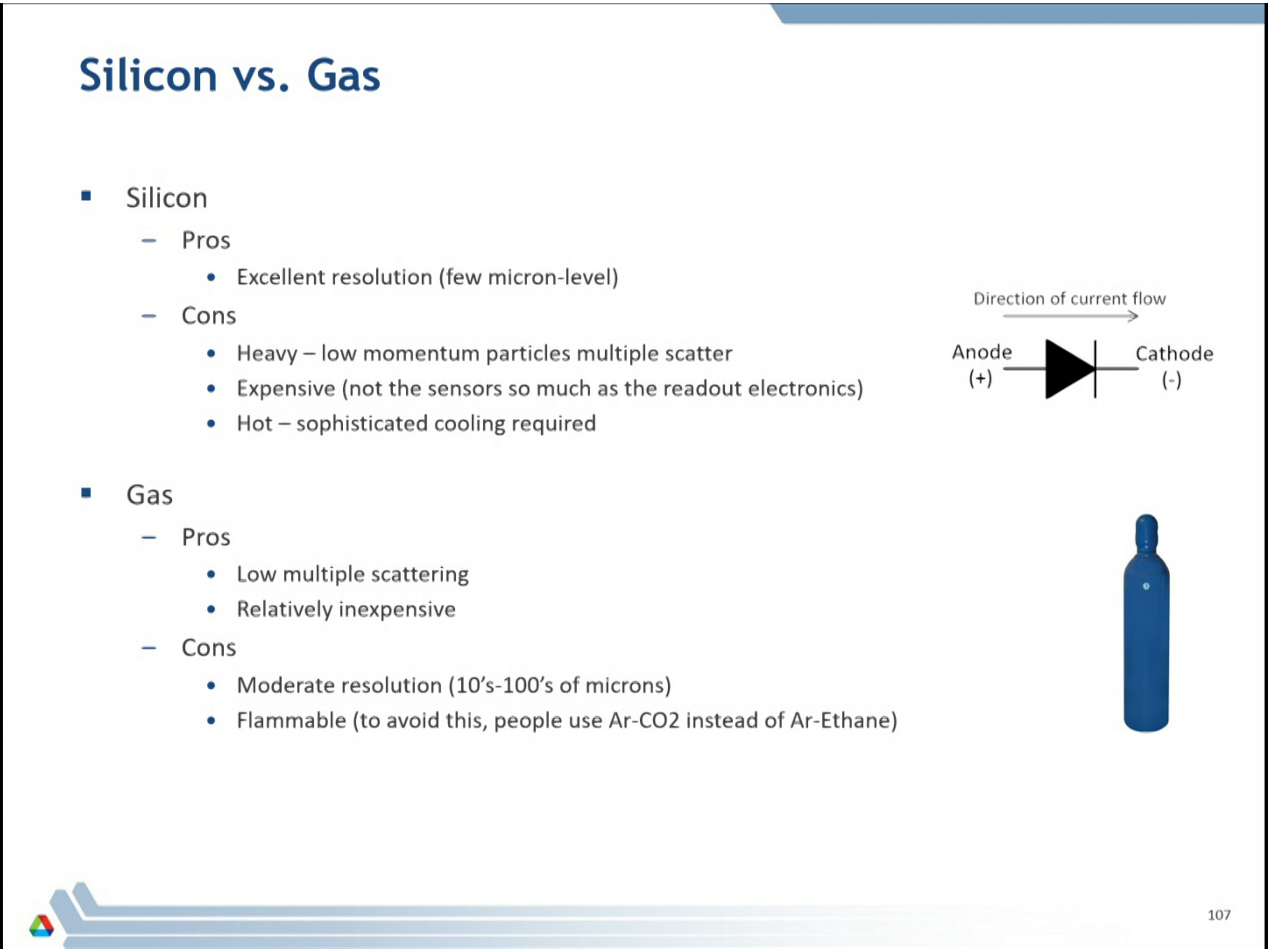
# How To Spot A Particle





- Really only one idea – as a particle traverses material, it ionizes that material

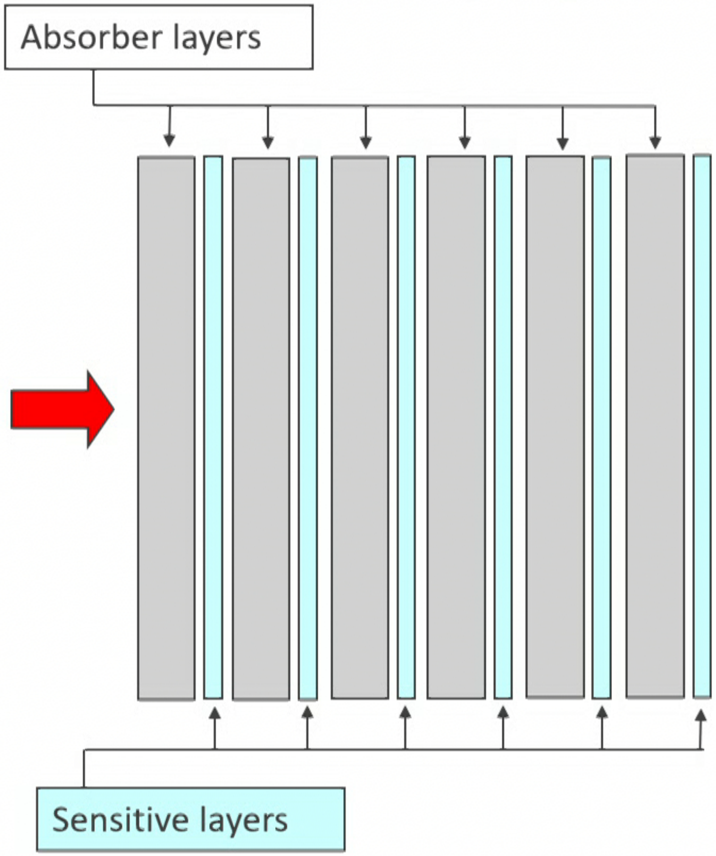
- Two common material choices
  - Silicon
    - Read out directly
  - Noble gasses, often with something flammable added to make things more interesting (e.g. Argon + Ethane)
    - Read out via wires or pads



# Silicon vs. Gas

- Silicon
  - Pros
    - Excellent resolution (few micron-level)
  - Cons
    - Heavy – low momentum particles multiple scatter
    - Expensive (not the sensors so much as the readout electronics)
    - Hot – sophisticated cooling required
- Gas
  - Pros
    - Low multiple scattering
    - Relatively inexpensive
  - Cons
    - Moderate resolution (10’s-100’s of microns)
    - Flammable (to avoid this, people use Ar-CO2 instead of Ar-Ethane)

# Fact Three: Sampling Calorimeters



Absorber layers

Sensitive layers

- A (constant) fraction of the incoming particle's energy gets converted to something that we can count (photons, electrons, etc...)

$$\frac{\delta E}{E} = \frac{\delta N}{N} \propto \frac{1}{\sqrt{N}} \propto \frac{1}{\sqrt{E}}$$

- In the approximation that each layer absorbs a constant fraction of the energy, the calorimeter depth grows logarithmically with energy.

As energy increases, calorimetry resolution improves.

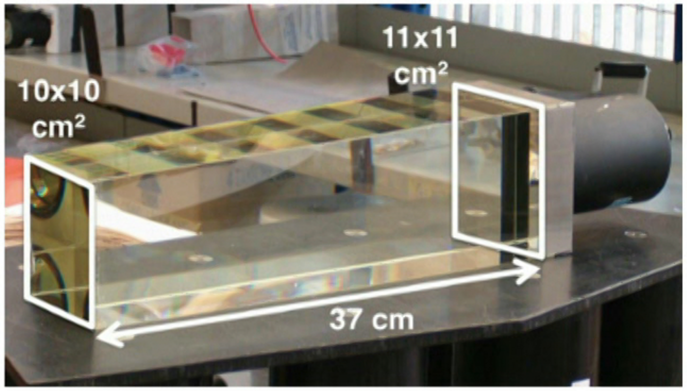
# Total Absorption Calorimeters

- Some calorimeters are homogeneous, i.e. they don't have separate absorber and sampler layers

- They convert the energy deposited into light, by either
  - Scintillation
  - Cerenkov radiation

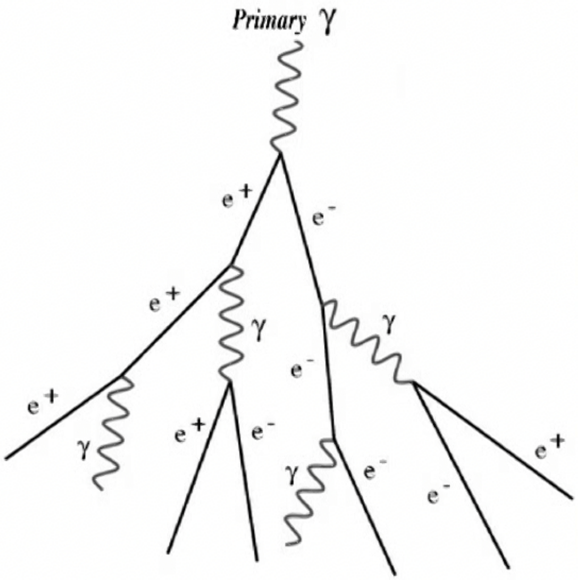
- Either way, it's still a statistical process so there is still √E behavior

- In general
  - These have better resolution
  - These are more expensive
  - Radiation damage can be an issue

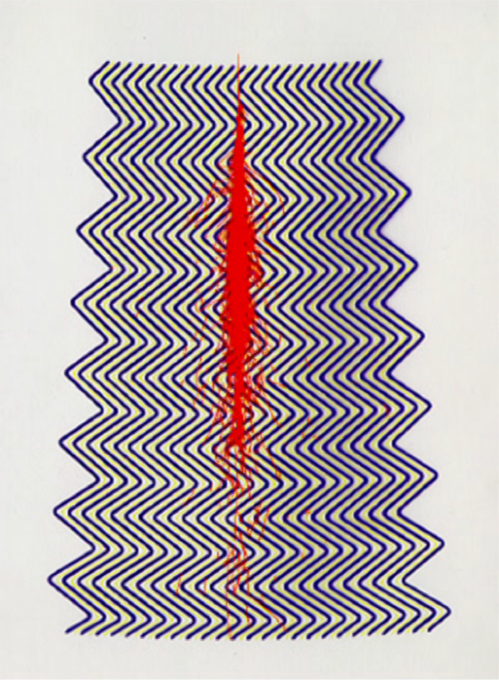


Block from OPAL. Uses Cerenkov radiation.

# Fact Three: Sampling Calorimeters (II)



A schematic of an electromagnetic shower



A GEANT simulation of an electromagnetic shower

Remember…

- EM showers all look the same

- Hadronic showers are like snowflakes
  - Every one is unique

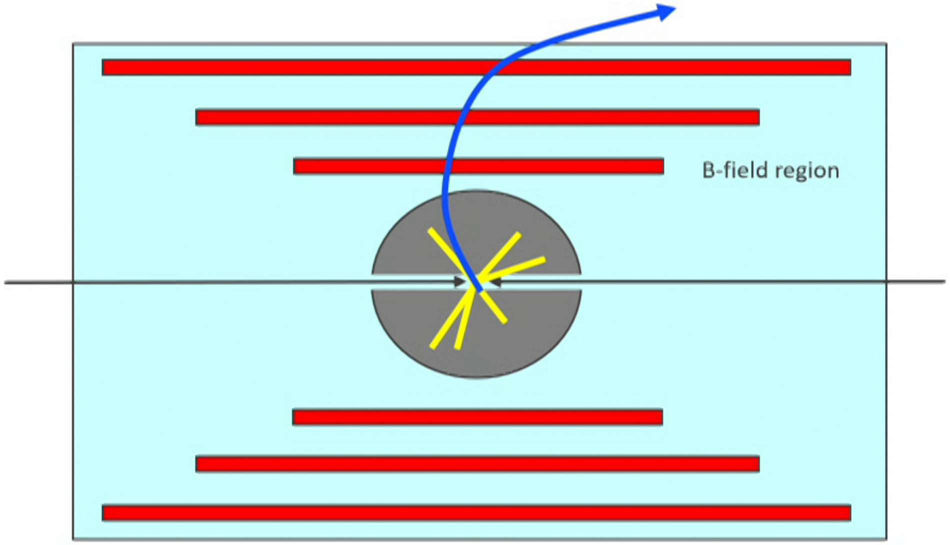
# Fact Four: Compromise is a fact of life

- Like it or not, experiments are constrained by resources
  - Every dollar that goes into one subsystem is a dollar that doesn't go into some other subsystem
  - Every inch that goes into one subsystem is an inch that doesn't go into some other subsystem
  - A collaboration with N members can't design an experiment that takes 2N members to build or operate.
  - Industrial production capacity is finite
    - Evidence: crystals, silicon wafers, liquid noble gasses

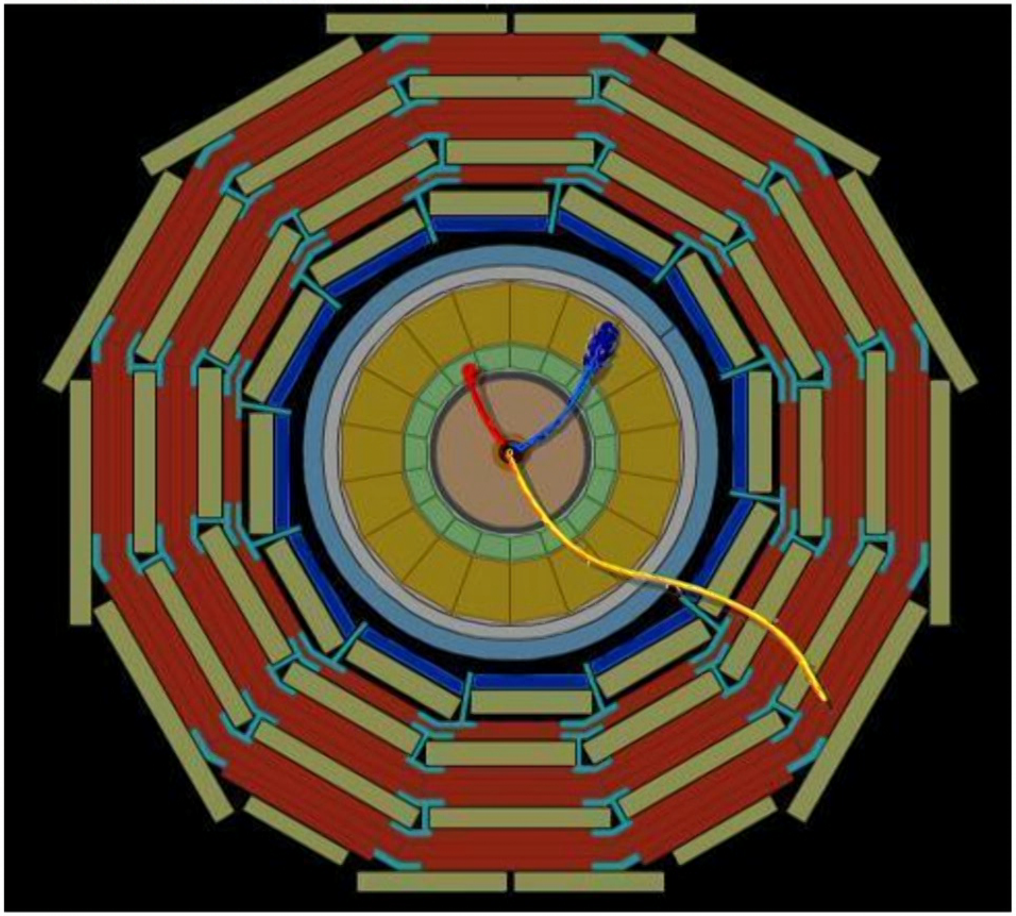
- Most experimenters have their own ideas on the best optimization
  - Individual interests and experience varies
  - The goal of a collaboration is to design a detector that everyone can live with – even if no single person thinks it's ideal.

# What Is The Iron Ball?

- Suppose you wanted a detector to look for a very heavy (800 GeV) Higgs via the decay H → ZZ followed by Z → μμ (both Z's)
  - This is a very rare process → increase the luminosity
  - Handle this increased luminosity by only looking at muons:

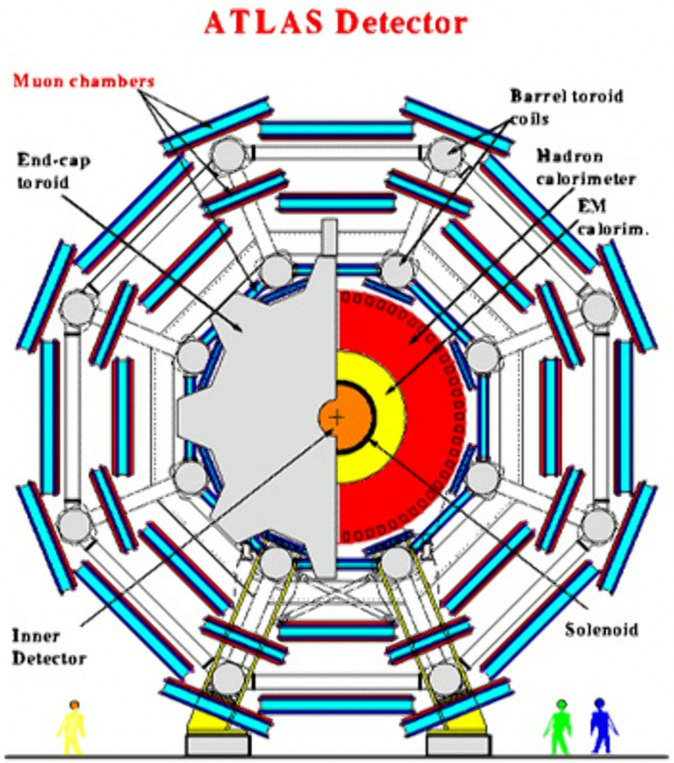


# CMS Muon Detectors

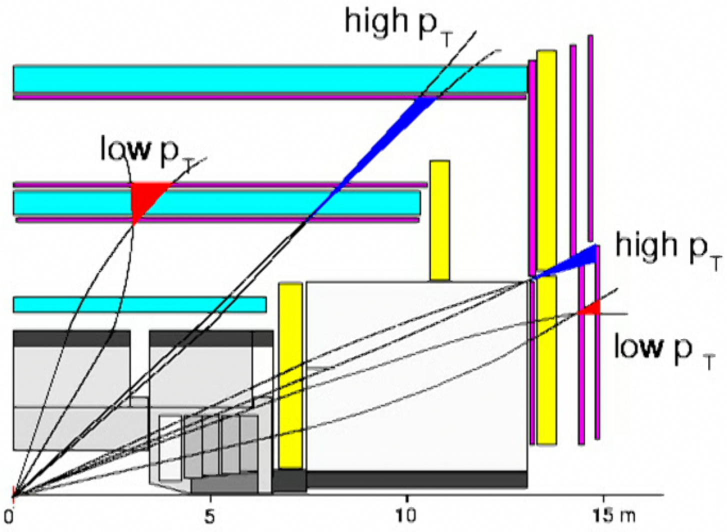


- CMS uses the return field of their central solenoid to measure muon momenta
- Four planes of detector stations inside the steel measure the muon’s tracks.
- Low $p_T$ muons range out in the steel, providing an additional measurement.

# The ATLAS Muon Spectrometer



Beam's eye view



How muon trajectories bend in the magnetic field of the toroids.

Energy stored in the magnetic field is ~1.2 GJ.

Energy stored in a lightning bolt is ~1.5 GJ.

Pictures from Jim Shank, Boston University

# Comparing Design Philosophies

Emphasizes

- CMS uses as their magnetic field the return field through the iron
  - Allows one to go to large fields...
  - ...which means small radii... (Figure of merit is $BL^2$)
  - ...which means that their calorimeter can be small (and expensive per unit volume).

$B$

- ATLAS uses air core toroids
  - Very good resolution at high momentum
    - Up to ~TeV scale muons
  - Requires a lot of space

$L^2$

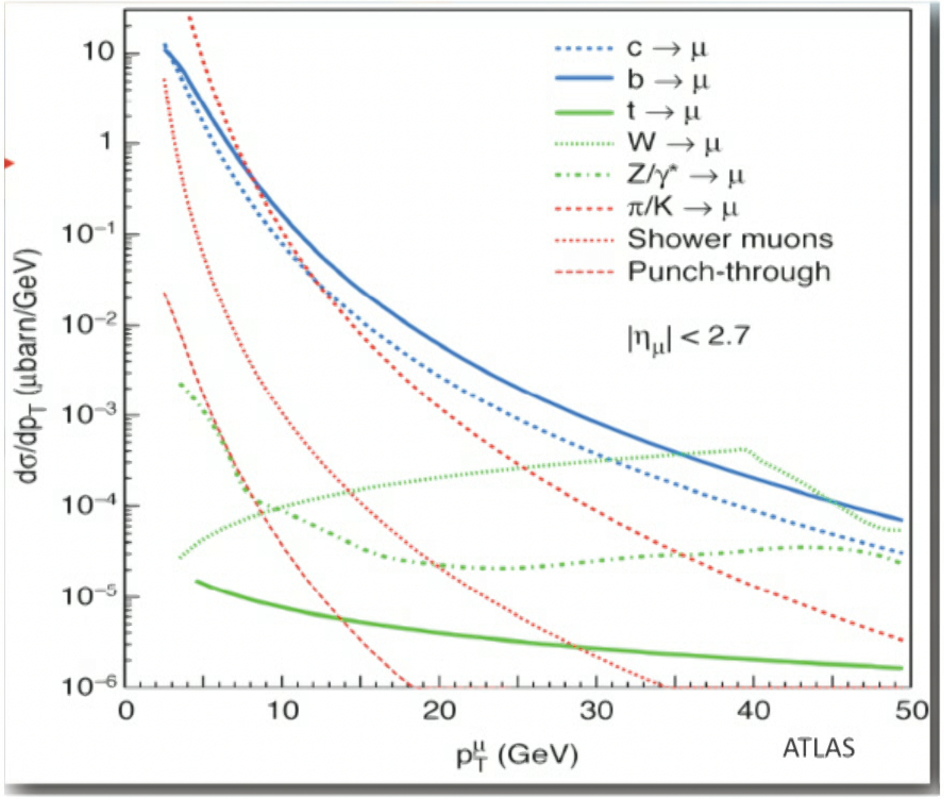
Neither experiment is an iron ball, but elements of the iron ball design influenced both detectors.

# Anticipated Muon Backgrounds

- The main background to muons is other muons

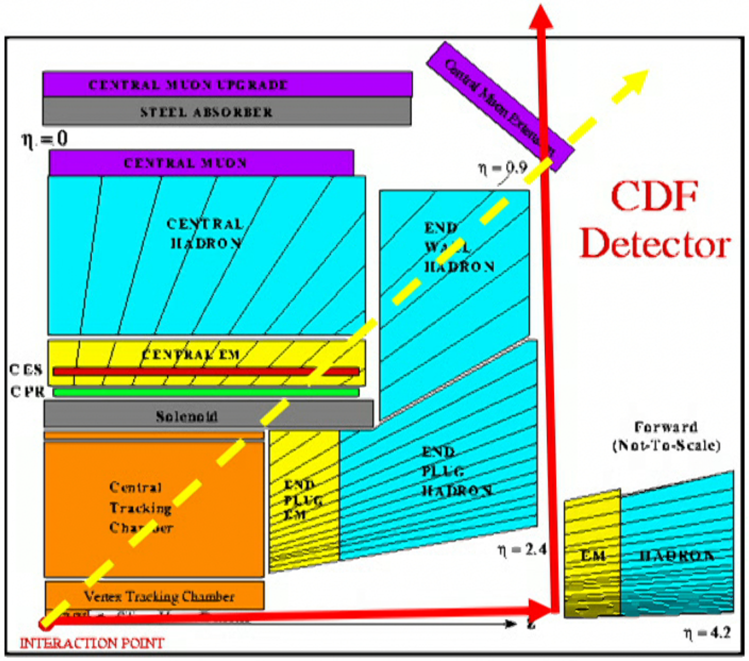
- What I mean is the main background to muons are muons from a source we are not particularly interested in.
  - Example: pion or kaon decays

- Improving purity beyond ATLAS or CMS hardly helps.



This is simulation – the equivalent plot for data exists, but is optimized to make a different point.

# Unanticipated Muon Backgrounds



The “CMX Ricochet”

- There is no such thing as a muon detector:
  - They are charged particle detectors, behind lots of steel
  - If there is any path around the steel, particles will find it
  - Since there are a million hadrons per muon, even if this is unlikely, it can be an important background
- At the LHC, the cavern backgrounds (secondary particles) might be large.
  - The experiments have to worry about particles “raining down” into their detectors.
  - This is VERY difficult to predict from first principles. It has to be measured. Fortunately, it turned out not to be a problem.
  - This time.

# Dimuon Mass

Requiring two muons is enough by itself to give a clean Z signal – and many other particles!

The background is real muons – just from other sources.