FROM RAW DATA TO PHYSICS

ANNA SFYRLA (UNIV. OF GENEVA)

SUMMER STUDENT LECTURES

JULY 2016

CONTENTS

Lecture 1

- **RAW data to Physics – step by step**
  - What does it take from getting the data out of the detector to producing a physics result.

Lecture 2

- **From RAW data to Standard Model Particles**
  - about measuring the properties of the ‘final’ particles created from a proton-proton interaction.

Lecture 3

- **From Standard Model Particles to measurements and searches**
  - about how…

ASSUMPTIONS

- You have never done a physics analysis.
- You know a bit about the LHC.
- You know a bit about a multi-purpose high-energy-physics detector.
- You know a bit about how we get to RAW data.

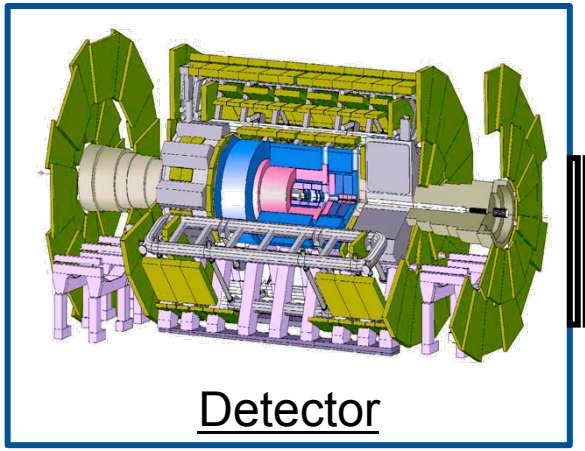
DISCLAIMER

- These lectures will have a “slight” bias towards ATLAS.

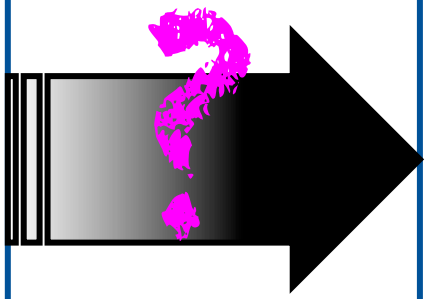
FROM RAW DATA TO PHYSICS

LECTURE 1

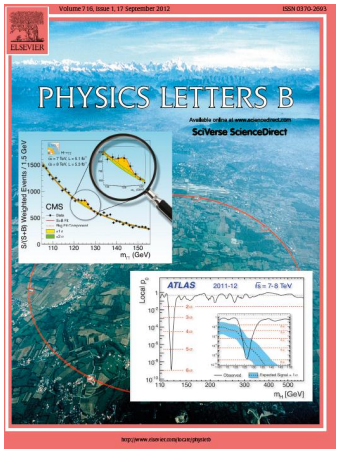
How do we deal with physics events from when they leave the detector till when they make it into our publications?

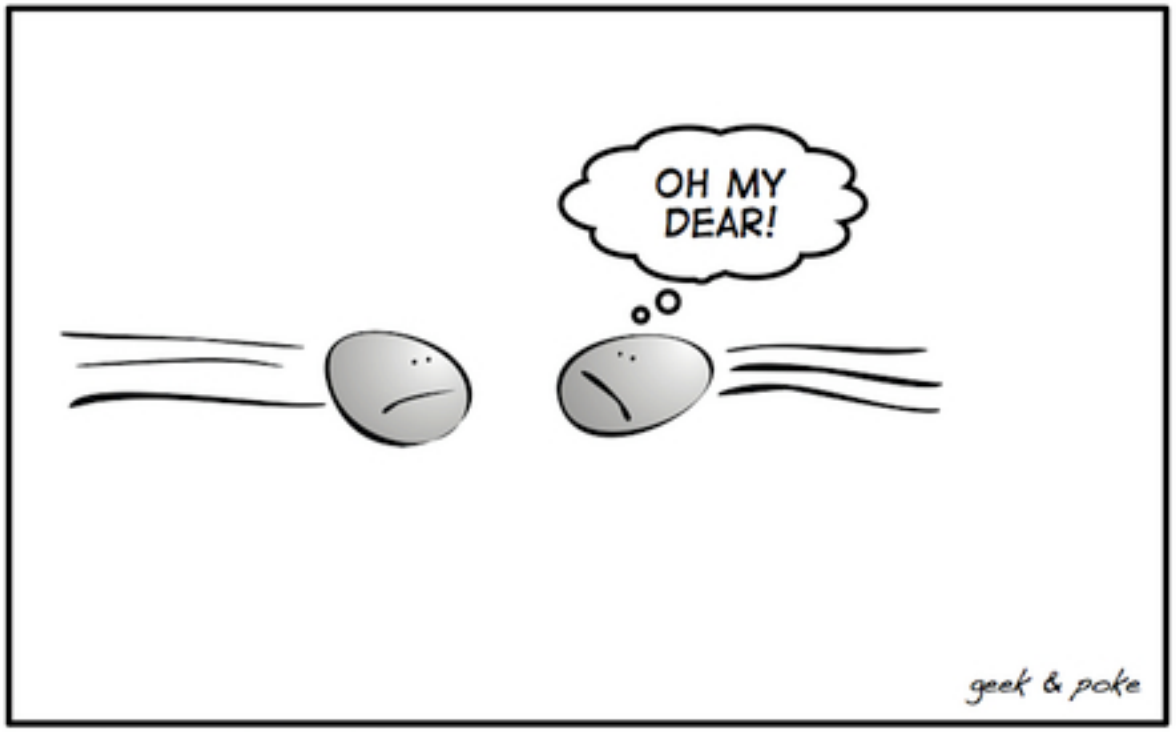


Detector



Publication

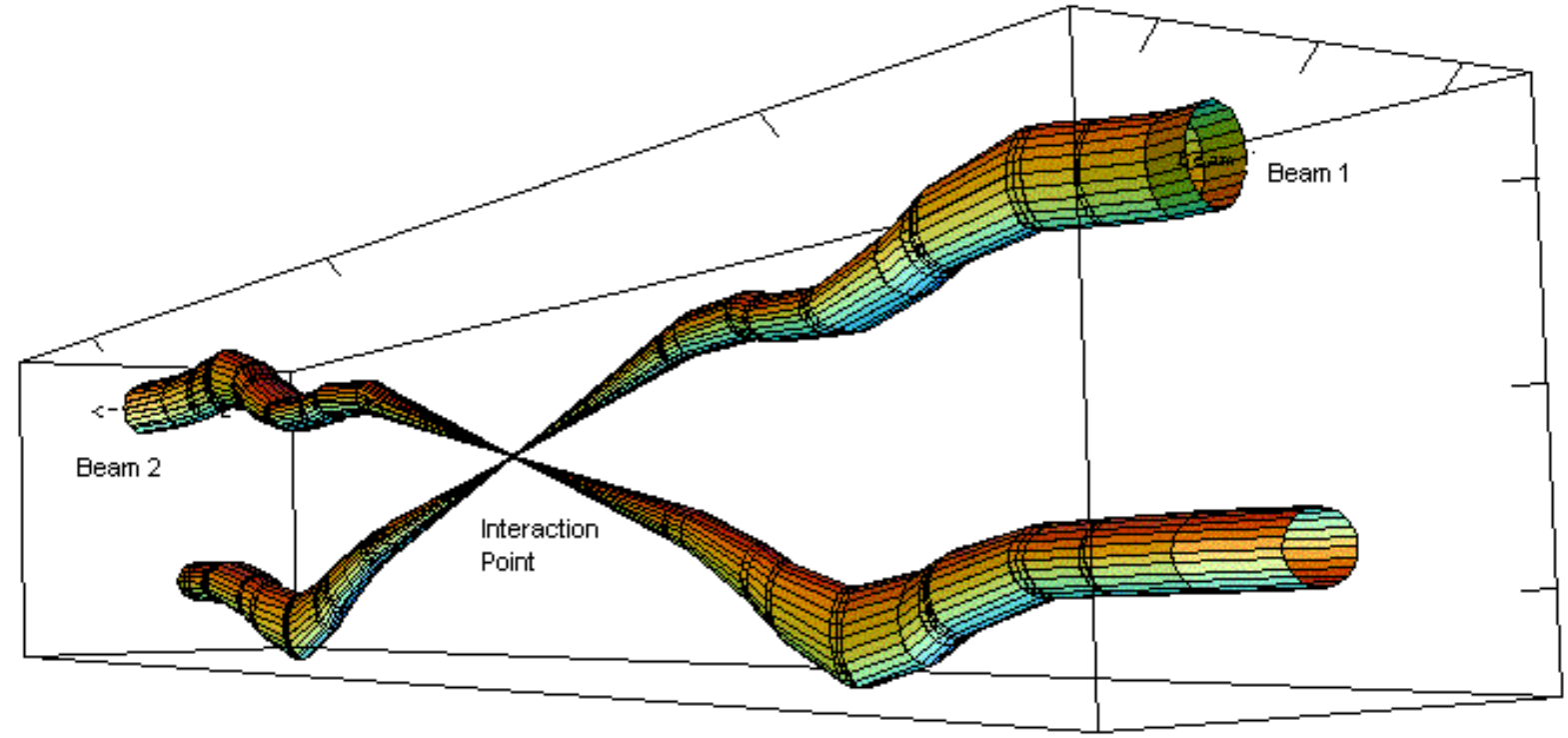




LATELY INSIDE THE LHC:
2 PROTONS 0.00000000000000000001 SEC BEFORE THE COLLISION

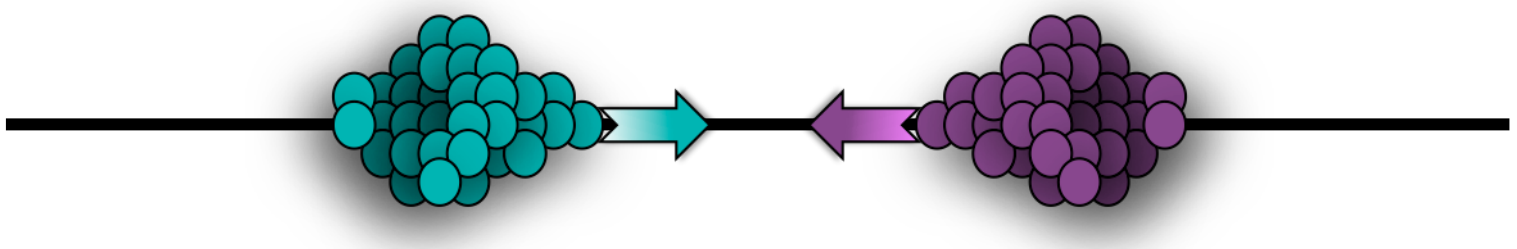
WHAT IS AN EVENT?

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



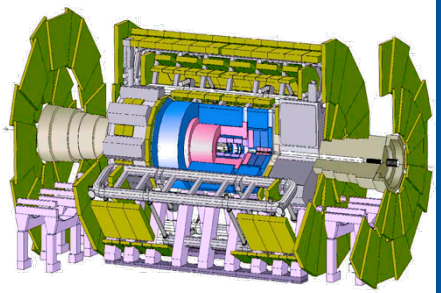
Relative beam sizes around IP1 (Atlas) in collision

WHAT IS AN EVENT?



Proton bunches
>10^{11} protons/bunch
colliding at **13TeV** and at **40MHz in Run-2**
collided at **7/8TeV** and at **20MHz in Run-1**

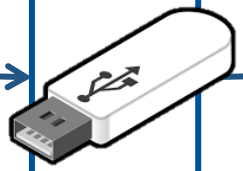
AN EVENT'S LIFETIME



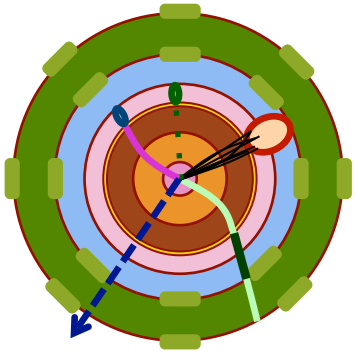
Detector

Trigger

Data Prep



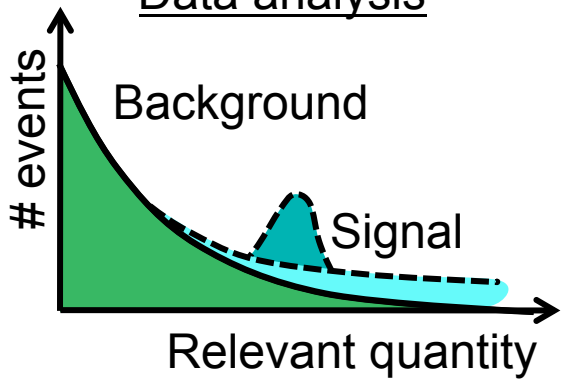
Reconstruction



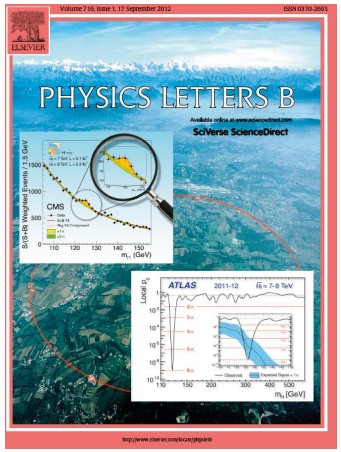
Calibration

Monte Carlo Simulations

Data analysis

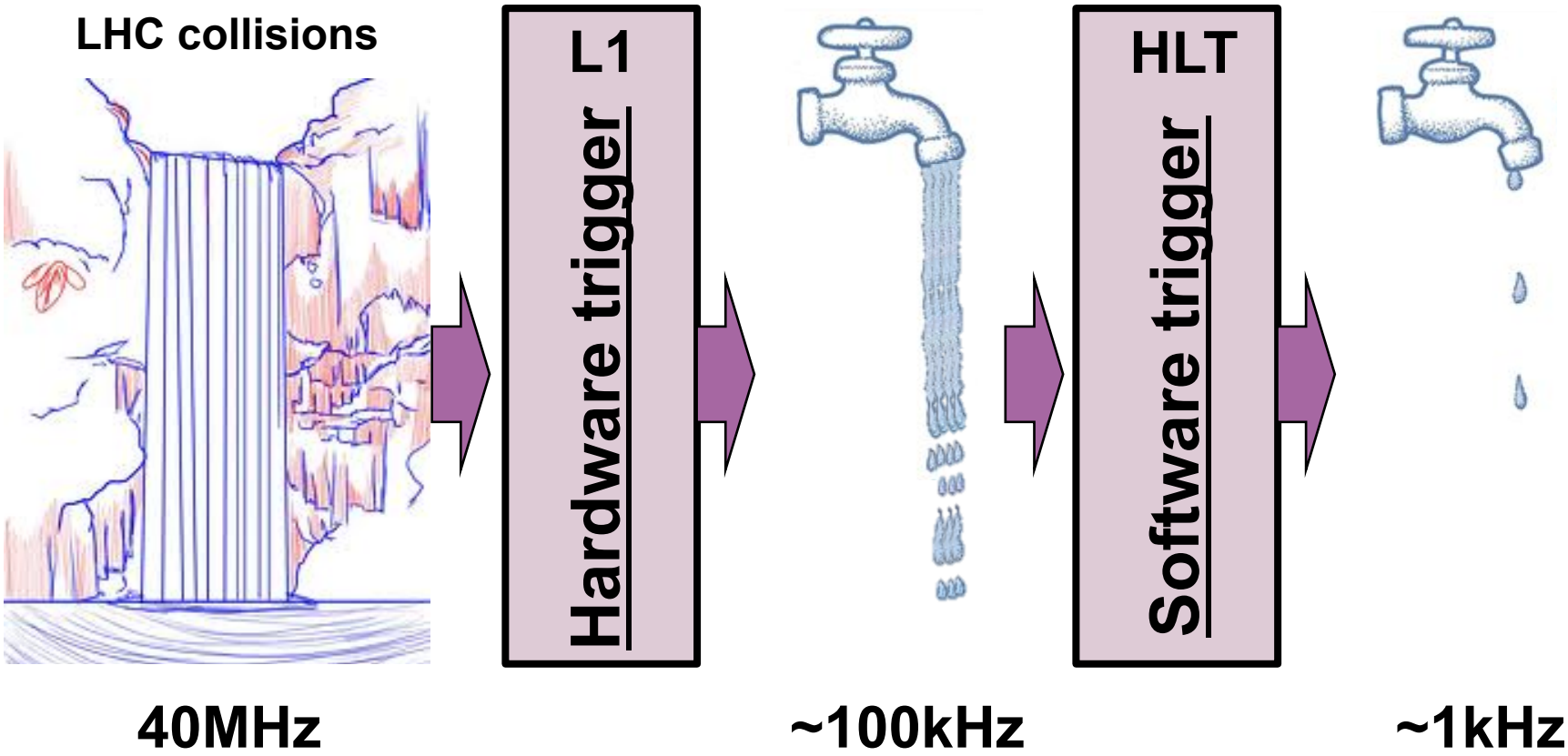


Publication



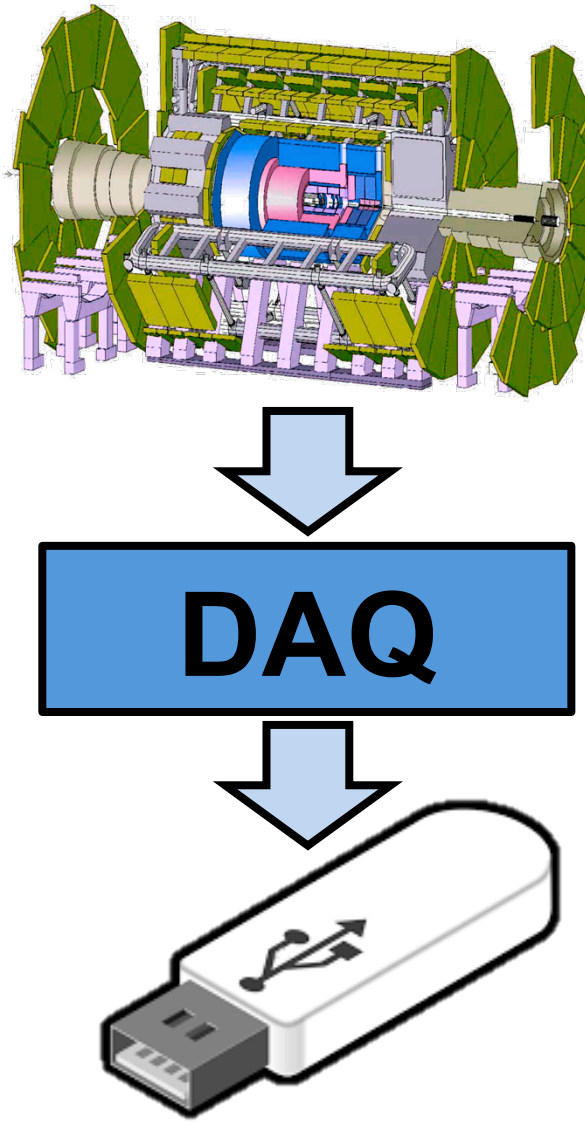
TRIGGERING ON PHYSICS

The ATLAS / CMS paradigm



THE DATA ACQUISITION

At every trigger accept:



WHAT DOES RAW CONTAIN?

A simple example from the trigger on ATLAS (run1 data)

0x00000015	0x20000e3f	536874559	lvl1 trigger info[0]
0x00000016	0x100000c0	268435648	lvl1 trigger info[1]
0x00000017	0x8000043f	2147484735	lvl1 trigger info[2]
0x00000018	0x00021007	135175	lvl1 trigger info[3]
0x00000019	0x00000e10	3600	lvl1 trigger info[4]
0x0000001a	0x00080000	524288	lvl1 trigger info[5]
0x0000001b	0x02c00400	46138368	lvl1 trigger info[6]
0x0000001c	0x00020001	131073	lvl1 trigger info[7]
0x0000001d	0x00000816	2070	lvl1 trigger info[8]
0x0000001e	0x100000c0	268435648	lvl1 trigger info[9]
0x0000001f	0x80000018	2147483672	lvl1 trigger info[10]
0x00000020	0x00021001	135169	lvl1 trigger info[11]
0x00000021	0x00000e10	3600	lvl1 trigger info[12]
0x00000022	0x00000000	0	lvl1 trigger info[13]
0x00000023	0x02c00400	46138368	lvl1 trigger info[14]
0x00000024	0x00020000	131072	lvl1 trigger info[15]
0x00000025	0x00000010	16	lvl1 trigger info[16]
0x00000026	0x00000000	0	lvl1 trigger info[17]
0x00000027	0x00000008	8	lvl1 trigger info[18]
0x00000028	0x00000000	0	lvl1 trigger info[19]
0x00000029	0x00000810	2064	lvl1 trigger info[20]
0x0000002a	0x00000000	0	lvl1 trigger info[21]
0x0000002b	0x00000400	1024	lvl1 trigger info[22]
0x0000002c	0x00000000	0	lvl1 trigger info[23]

L1 Trigger Bits Before Prescale

L1 Trigger Bits After Prescale

L1 Trigger Bits After Veto

WHAT DOES RAW CONTAIN?

A simple example from the trigger on ATLAS (run1 data)

0x00000015	0x20000e3f	536874559	lvl1 trigger info[0]
0x00000016	0x100000c0	268435648	lvl1 trigger info[1]
0x00000017	0x8000043f	2147484735	lvl1 trigger info[2]
0x00000018	0x00021007	135175	lvl1 trigger info[3]
0x00000019	0x00000e10	3600	lvl1 trigger info[4]
0x0000001a	0x00080000	524288	lvl1 trigger info[5]
0x0000001b	0x02c00400	46138368	lvl1 trigger info[6]
0x0000001c	0x00020001	131073	lvl1 trigger info[7]
0x0000001d	0x00000816	2070	lvl1 trigger info[8]
0x0000001e	0x100000c0	268435648	lvl1 trigger info[9]
0x0000001f	0x80000018	2147483672	lvl1 trigger info[10]
0x00000020	0x00021001	135169	lvl1 trigger info[11]
0x00000021	0x00000e10	3600	lvl1 trigger info[12]
0x00000022	0x00000000	0	lvl1 trigger info[13]
0x00000023	0x02c00400	46138368	lvl1 trigger info[14]
0x00000024	0x00020000	131072	lvl1 trigger info[15]
0x00000025	0x00000010	16	lvl1 trigger info[16]
0x00000026	0x00000000	0	lvl1 trigger info[17]
0x00000027	0x00000008	8	lvl1 trigger info[18]
0x00000028	0x00000000	0	lvl1 trigger info[19]
0x00000029	0x00000810	2064	lvl1 trigger info[20]
0x0000002a	0x00000000	0	lvl1 trigger info[21]
0x0000002b	0x00000400	1024	lvl1 trigger info[22]
0x0000002c	0x00000000	0	lvl1 trigger info[23]

Enabled items, ID:
0, 1, 2, 3, 4, 5, 9, 10, 11, 29, 38, 39, 60, 64, 65, 66, 67, 68, 69, 74, 95, 96, 97, 98, 108, 113, 132, 137, 138, 139, 179, 202, 214, 215, 217, 224, 241

Enabled items, ID:
1, 2, 4, 11, 38, 39, 60, 67, 68, 95, 96, 108, 113, 132, 137, 138, 139, 202, 214, 215, 217, 241

Enabled items, ID:
4, 67, 132, 139, 202

WHAT DOES RAW CONTAIN?

A simple example from the trigger on ATLAS (run1 data)

0x00000015	0x20000e3f	536874559	lvl1 trigger info[0]
0x00000016	0x100000c0	268435648	lvl1 trigger info[1]
0x00000017	0x8000043f	2147484735	lvl1 trigger info[2]
0x00000018	0x00021007	135175	lvl1 trigger info[3]
0x00000019	0x00000e10	3600	lvl1 trigger info[4]
0x0000001a	0x00080000	524288	lvl1 trigger info[5]
0x0000001b	0x02c00400	46138368	lvl1 trigger info[6]
0x0000001c	0x00020001	131073	lvl1 trigger info[7]
0x0000001d	0x00000816	2070	lvl1 trigger info[8]
0x0000001e	0x100000c0	268435648	lvl1 trigger info[9]
0x0000001f	0x80000018	2147483672	lvl1 trigger info[10]
0x00000020	0x00021001	135169	lvl1 trigger info[11]
0x00000021	0x00000e10	3600	lvl1 trigger info[12]
0x00000022	0x00000000	0	lvl1 trigger info[13]
0x00000023	0x02c00400	46138368	lvl1 trigger info[14]
0x00000024	0x00020000	131072	lvl1 trigger info[15]
0x00000025	0x00000010	16	lvl1 trigger info[16]
0x00000026	0x00000000	0	lvl1 trigger info[17]
0x00000027	0x00000008	8	lvl1 trigger info[18]
0x00000028	0x00000000	0	lvl1 trigger info[19]
0x00000029	0x00000810	2064	lvl1 trigger info[20]
0x0000002a	0x00000000	0	lvl1 trigger info[21]
0x0000002b	0x00000400	1024	lvl1 trigger info[22]
0x0000002c	0x00000000	0	lvl1 trigger info[23]

Enabled items, ID:
0, 1, 2, 3, 4, 5, 9, 10, 11, 29, 38, 39, 60, 64, 65, 66, 67, 68, 69, 74, 95, 96, 97, 98, 108, 113, 132, 137, 138, 139, 179, 202, 214, 215, 217, 224, 241

Enabled items, ID:
1, 2, 4, 11, 38, 39, 60, 67, 68, 95, 96, 108, 113, 132, 137, 138, 139, 202, 214, 215, 217, 241

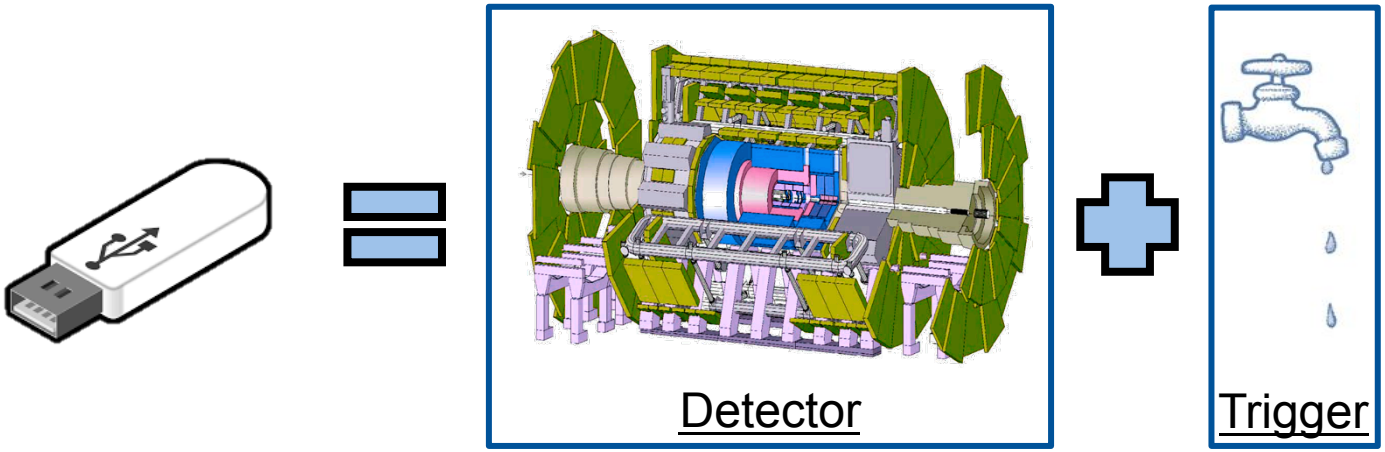
Enabled items, name:
L1_EM18VH,
L1_2TAU11I_EM14VH,
L1_2TAU11_TAU20_EM14VH,
L1_2TAU11I_TAU15,
L1_2EM6_EM16VH

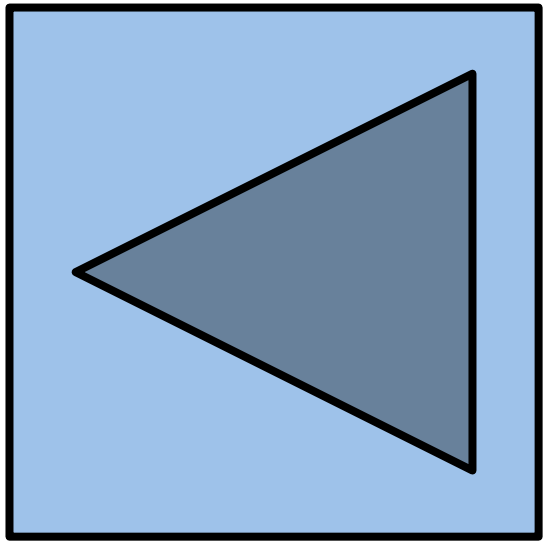
WHAT DOES RAW CONTAIN?

0x00000015	0x20000e3f	536874559	lvl1 trigger info[0]
0x00000016	0x100000c0	268435648	lvl1 trigger info[1]
0x00000017	0x8000043f	2147484735	lvl1 trigger info[2]
0x00000018	0x00021007	135175	lvl1 trigger info[3]
0x00000019	0x00000e10	3600	lvl1 trigger info[4]
0x0000001a	0x00080000	524288	lvl1 trigger info[5]
0x0000001b	0x02c00400	46138368	lvl1 trigger info[6]
0x0000001c	0x00020001	131073	lvl1 trigger info[7]
0x0000001d	0x00000816	2070	lvl1 trigger info[8]
0x0000001e	0x100000c0	268435648	lvl1 trigger info[9]
0x0000001f	0x80000018	2147483672	lvl1 trigger info[10]
0x00000020	0x00021001	135169	lvl1 trigger info[11]
0x00000021	0x00000e10	3600	lvl1 trigger info[12]
0x00000022	0x00000000	0	lvl1 trigger info[13]
0x00000023	0x02c00400	46138368	lvl1 trigger info[14]
0x00000024	0x00020000	131072	lvl1 trigger info[15]
0x00000025	0x00000010	16	lvl1 trigger info[16]
0x00000026	0x00000000	0	lvl1 trigger info[17]
0x00000027	0x00000008	8	lvl1 trigger info[18]
0x00000028	0x00000000	0	lvl1 trigger info[19]
0x00000029	0x00000810	2064	lvl1 trigger info[20]
0x0000002a	0x00000000	0	lvl1 trigger info[21]
0x0000002b	0x00000400	1024	lvl1 trigger info[22]
0x0000002c	0x00000000	0	lvl1 trigger info[23]

- 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!!**

WHAT DOES RAW CONTAIN?





THE LARGE HADRON COLLIDER

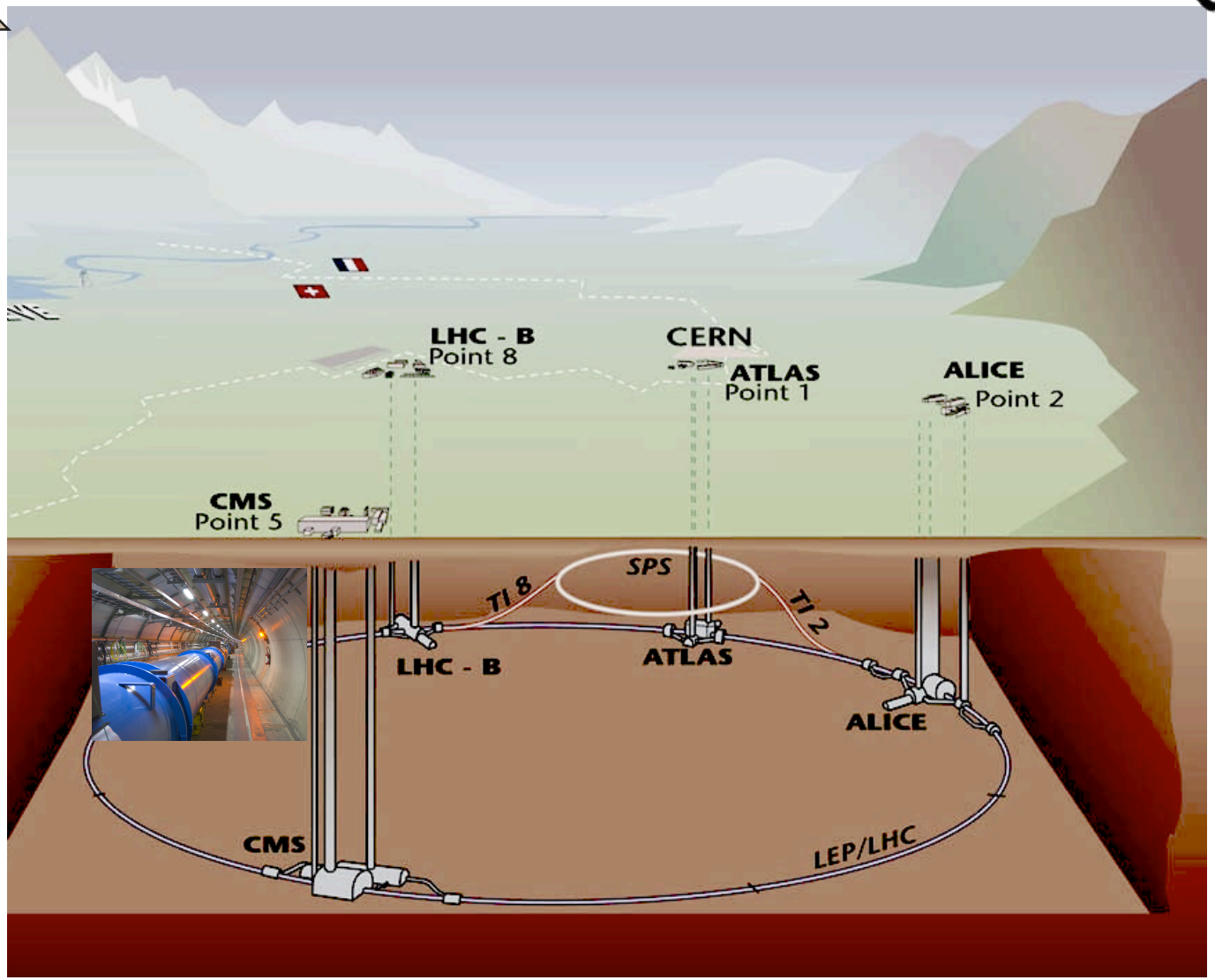
New Particle!
(<<**mHz?**)

p-p collisions with interesting parton interactions
(<kHz)

p p

~25 p-p collisions/bc

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



THE LARGE HADRON COLLIDER

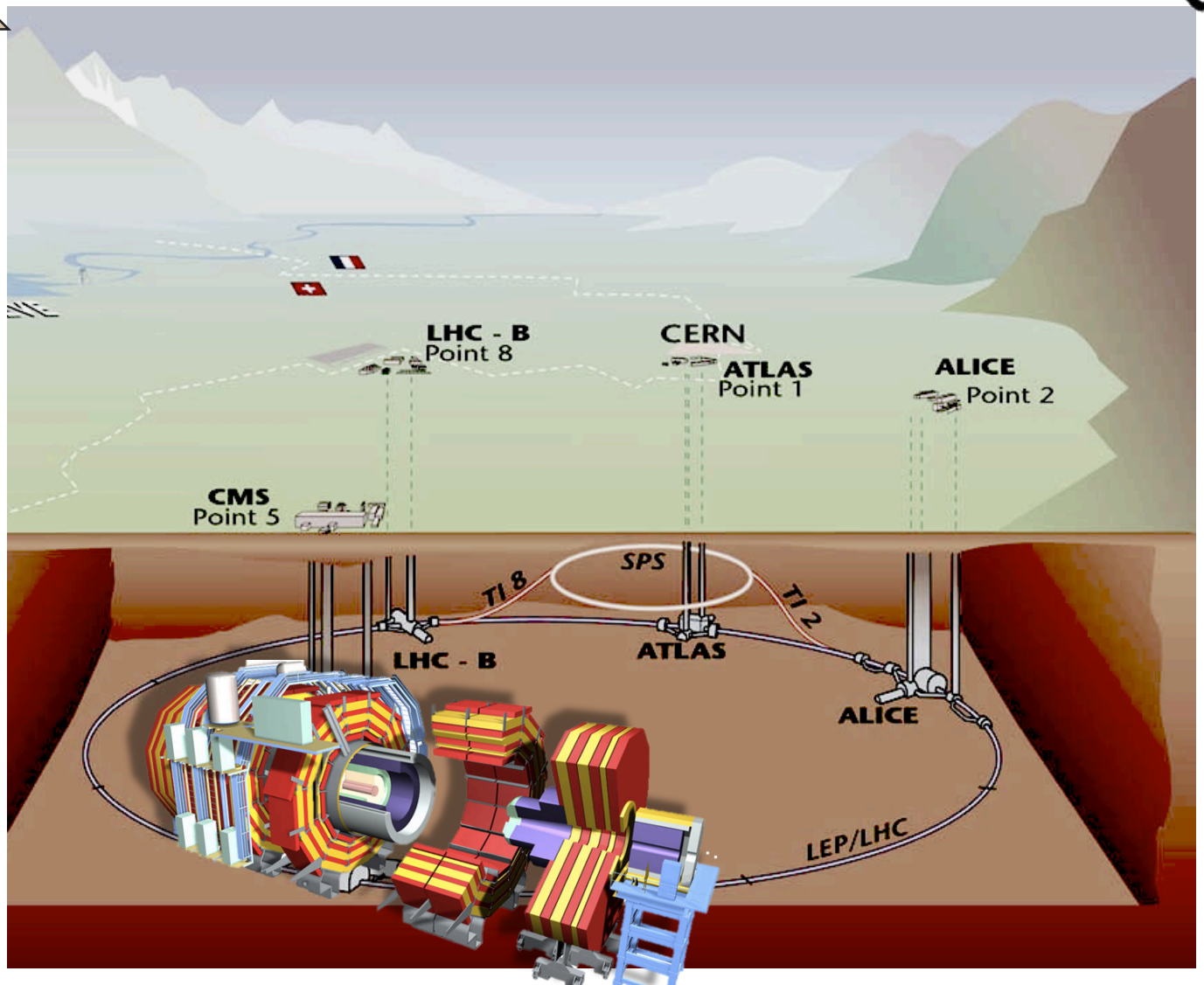
New Particle!
(<<**mHz?**)

p-p collisions with interesting parton interactions
(<kHz)

p p

~25 p-p collisions/bc

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



THE LARGE HADRON COLLIDER

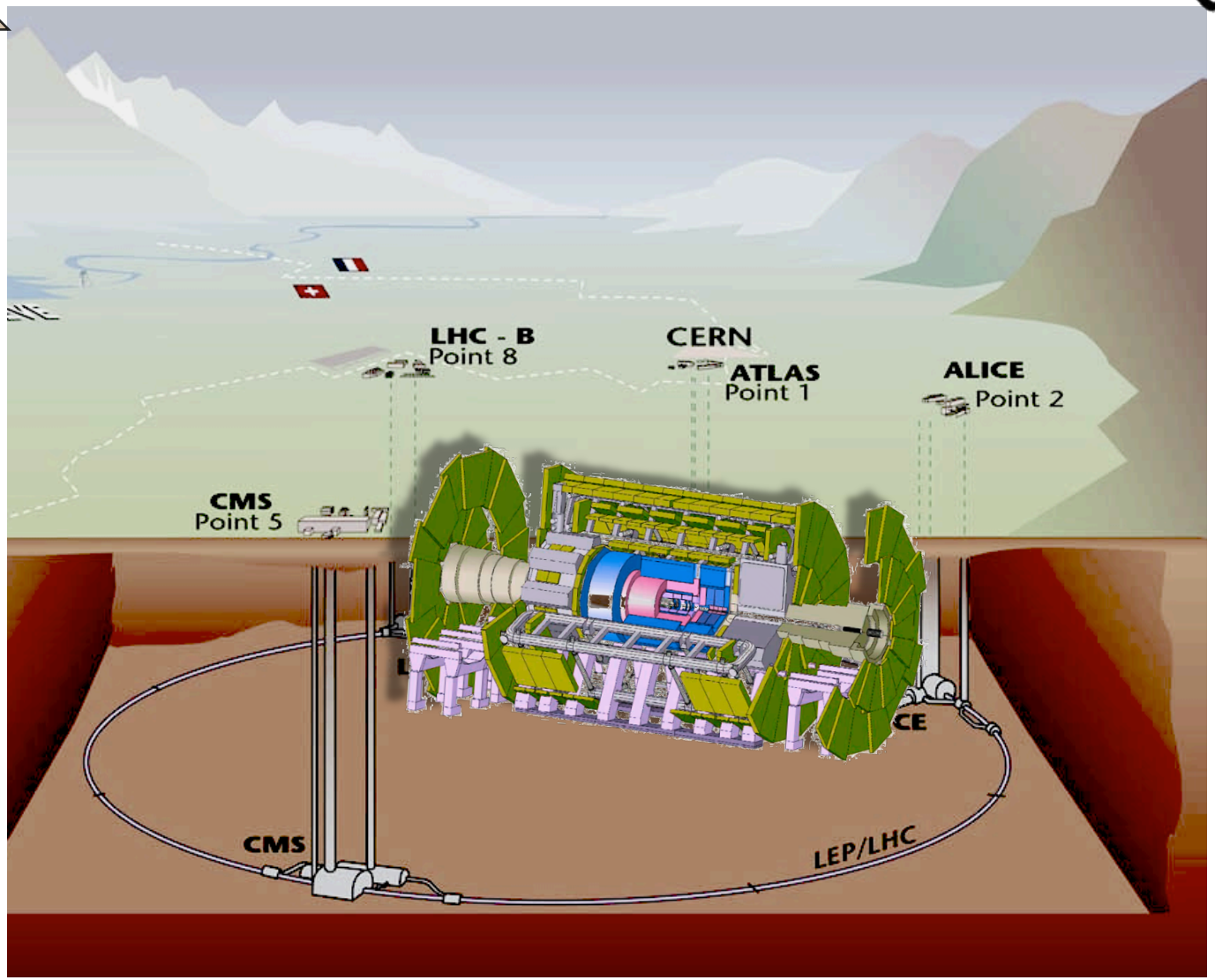
New Particle!
(<<**mHz?**)

p-p collisions with interesting parton interactions
(<kHz)

p p

~25 p-p collisions/bc

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



THE LARGE HADRON COLLIDER

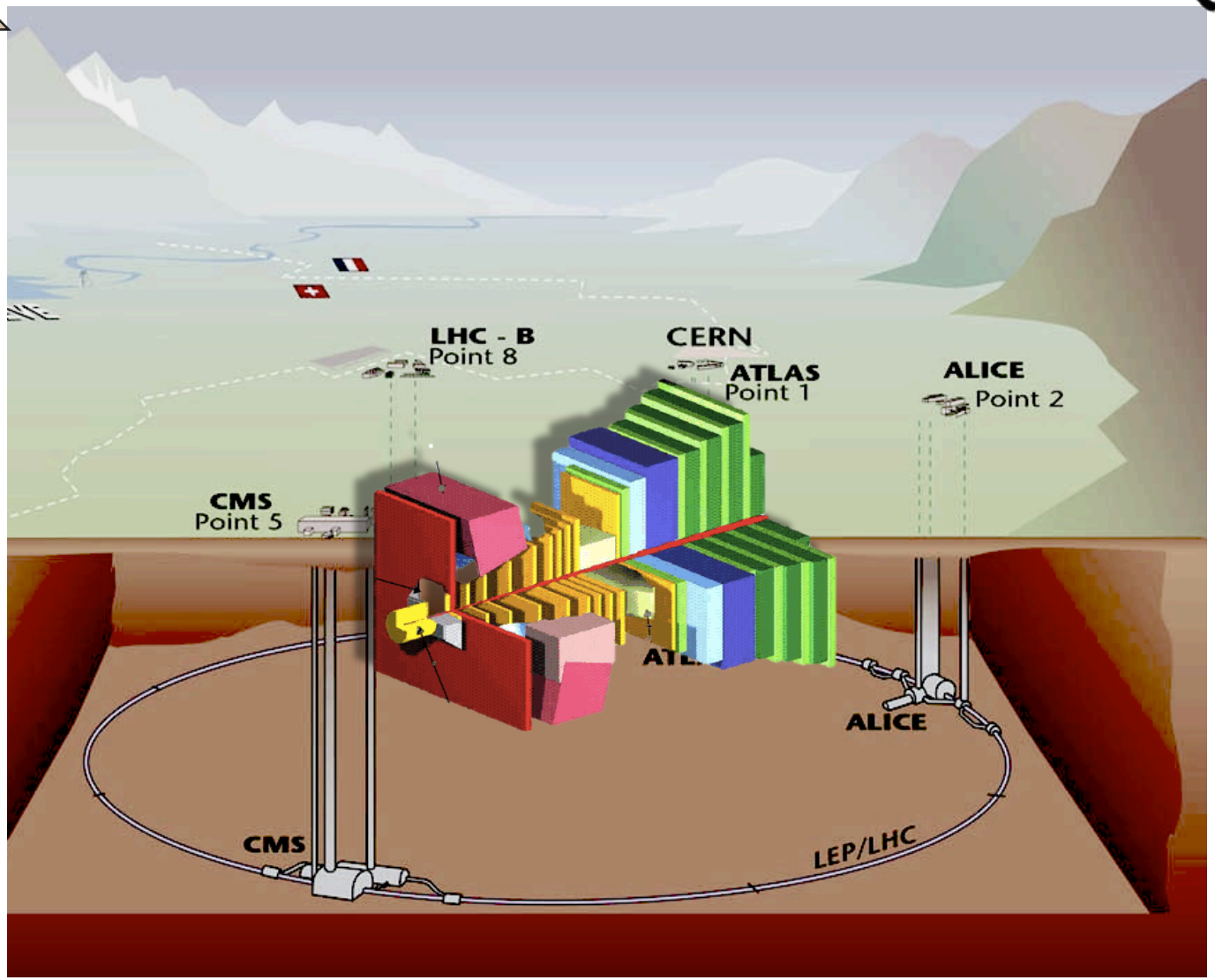
New Particle!
(<<**mHz?**)

p-p collisions with interesting parton interactions
(<kHz)

p p

~25 p-p collisions/bc

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



THE LARGE HADRON COLLIDER

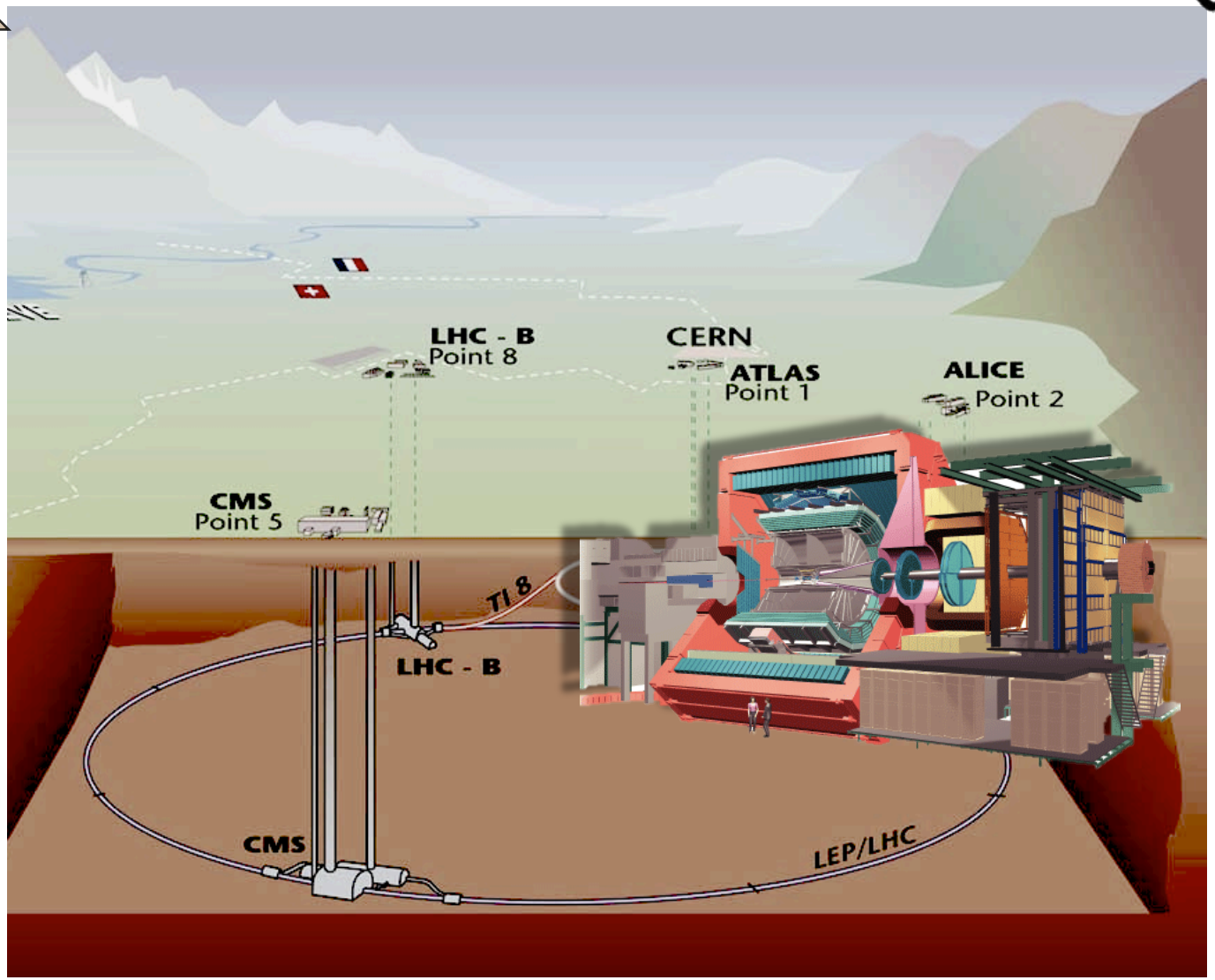
New Particle!
(<<**mHz?**)

p-p collisions with interesting parton interactions
(<kHz)

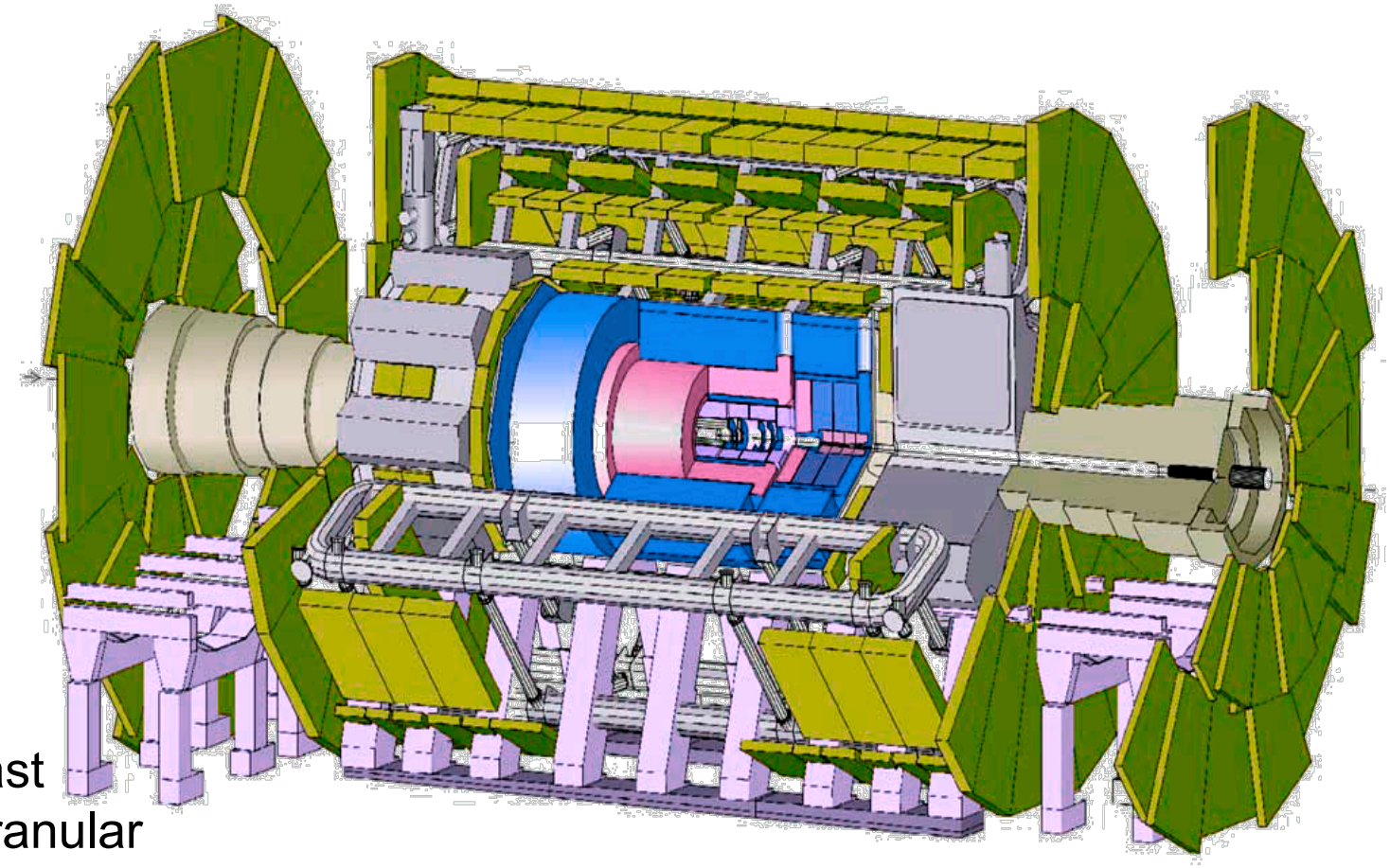
p p

~25 p-p collisions/bc

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

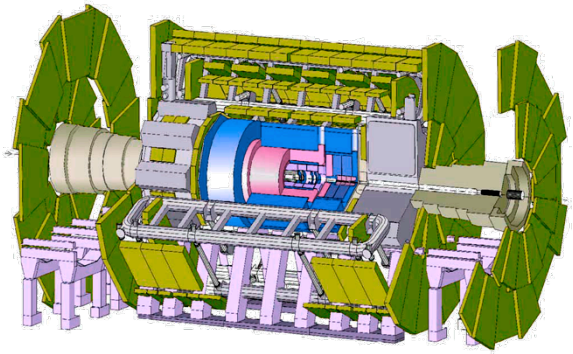


A DETECTOR (E.G. ATLAS)

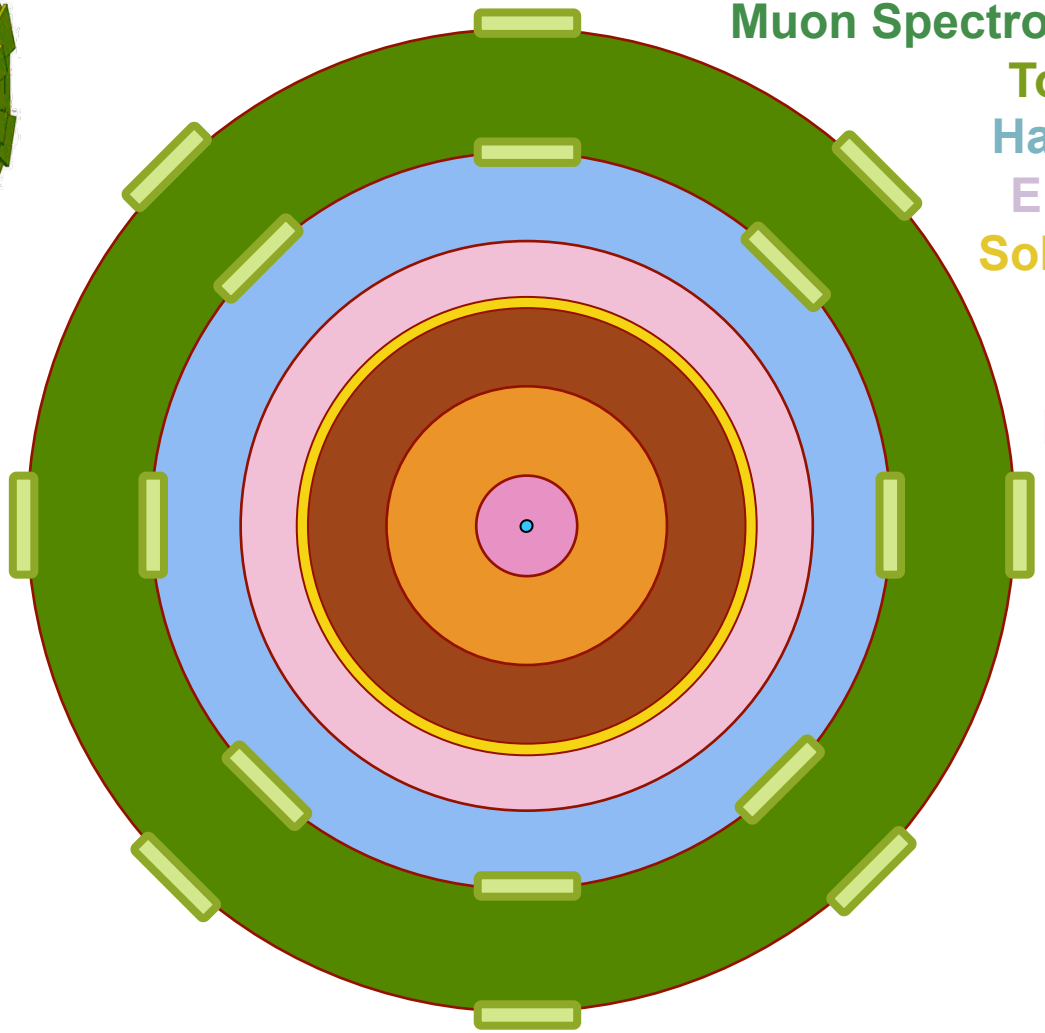


- ✓ Fast
- ✓ Granular
- ✓ Resistant to radiation

A DETECTOR (E.G. ATLAS)



Simplified Detector Transverse View

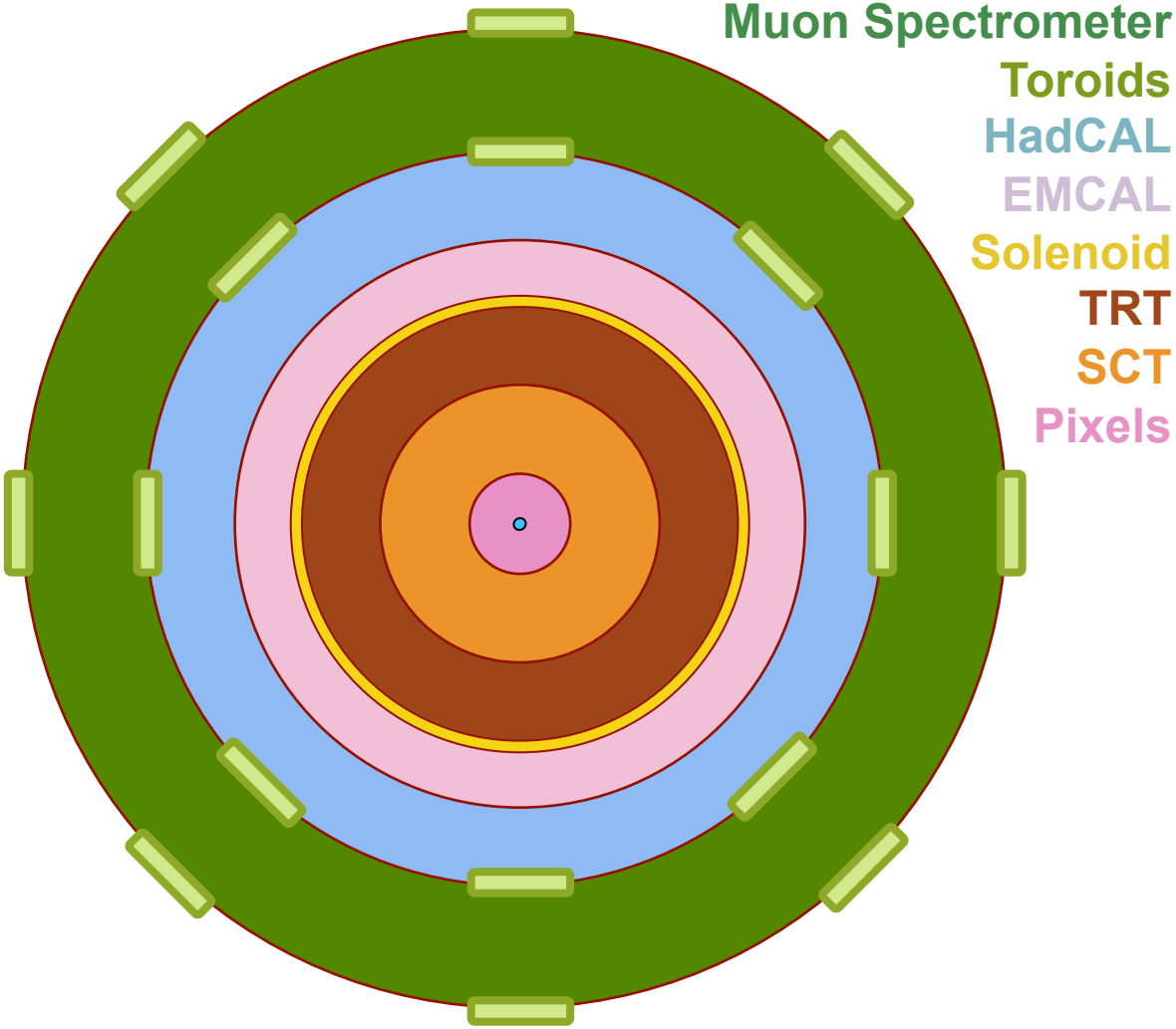


A DETECTOR (E.G. ATLAS)

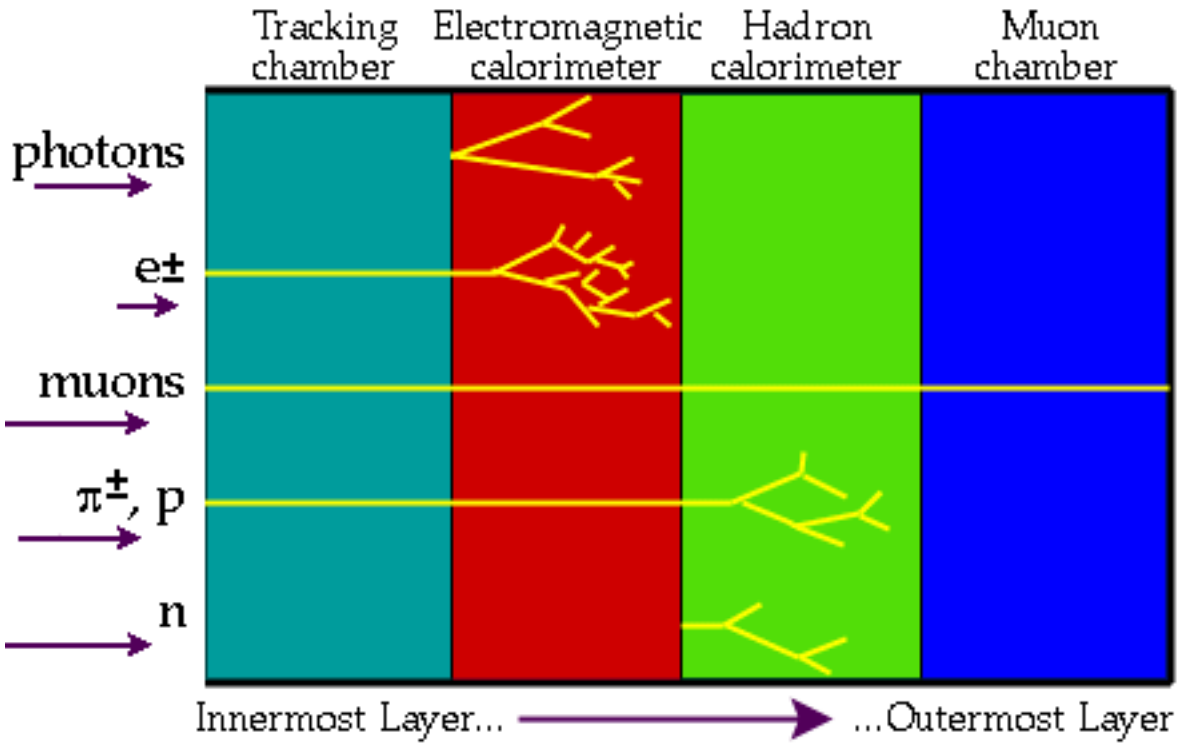
	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2.2 eV ν_e	<0.2 MeV ν_μ	<16 MeV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

Bosons

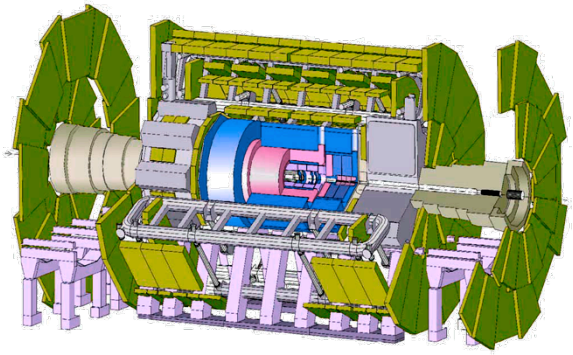
Simplified Detector Transverse View



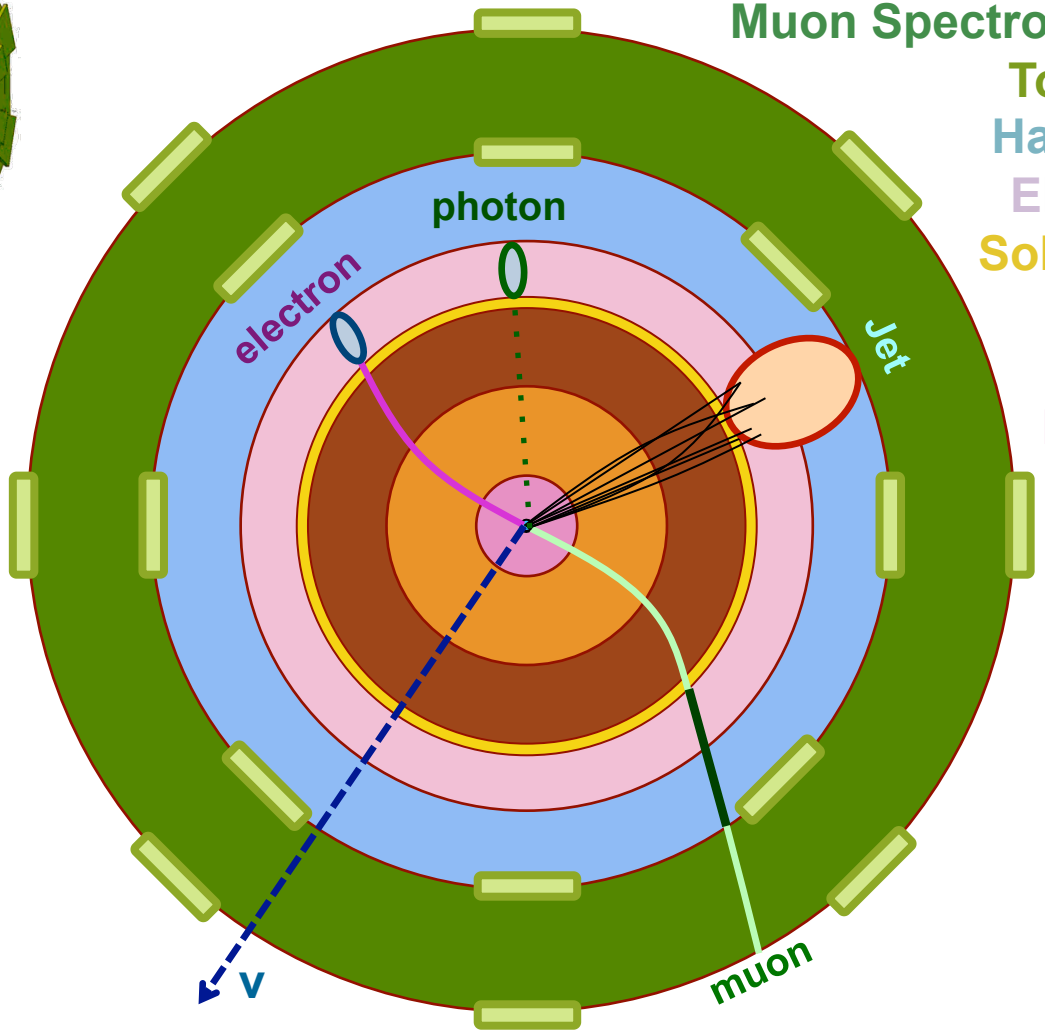
PARTICLES THROUGH MATTER



A DETECTOR (E.G. ATLAS)



Simplified Detector Transverse View



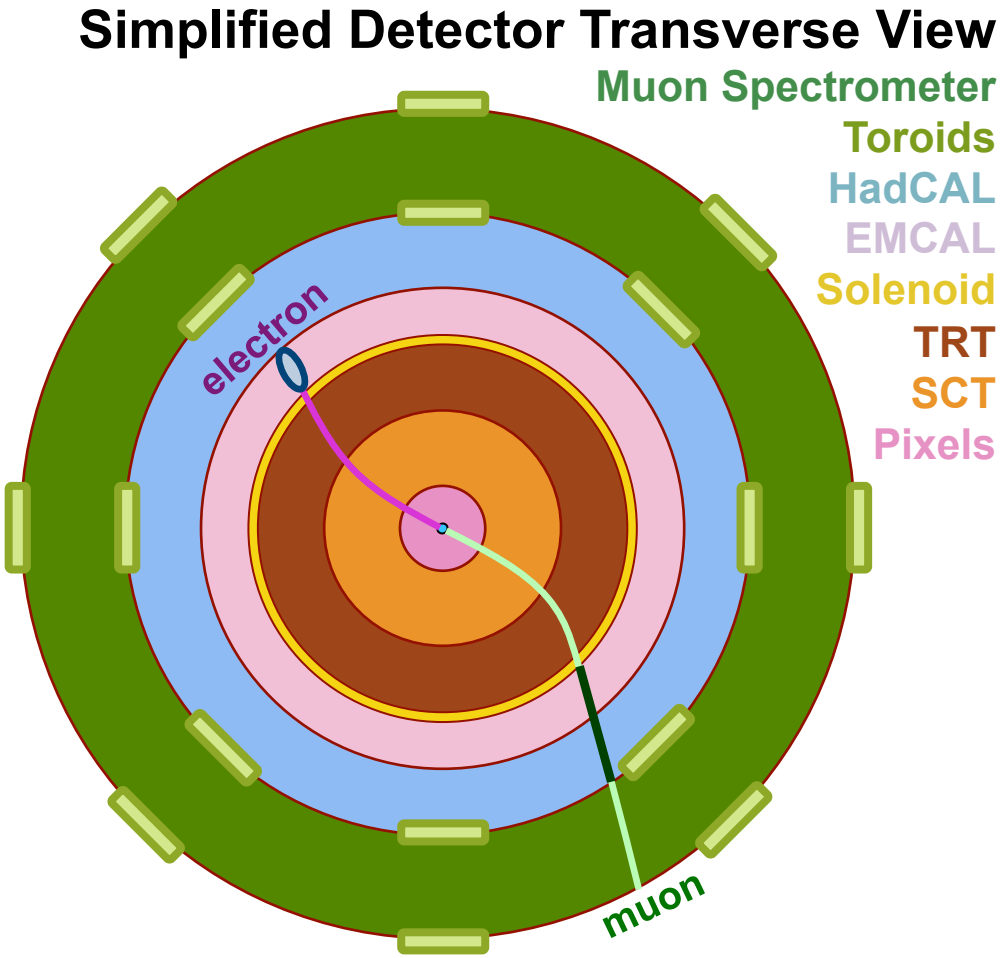
	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2.2 eV ν_e	<0.2 MeV ν_μ	<16 MeV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

Bosons

RECONSTRUCTING PARTICLES

	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2 eV ν_e	<2 eV ν_μ	<2 eV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

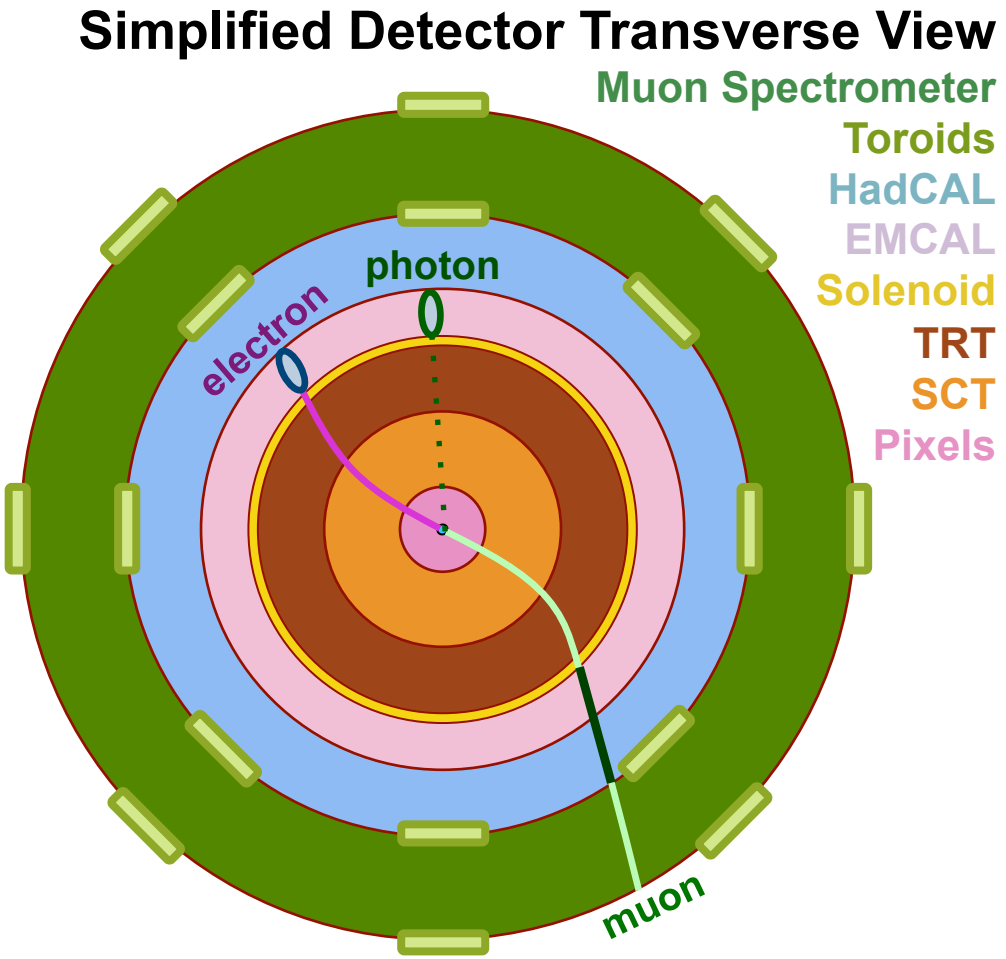
Bosons



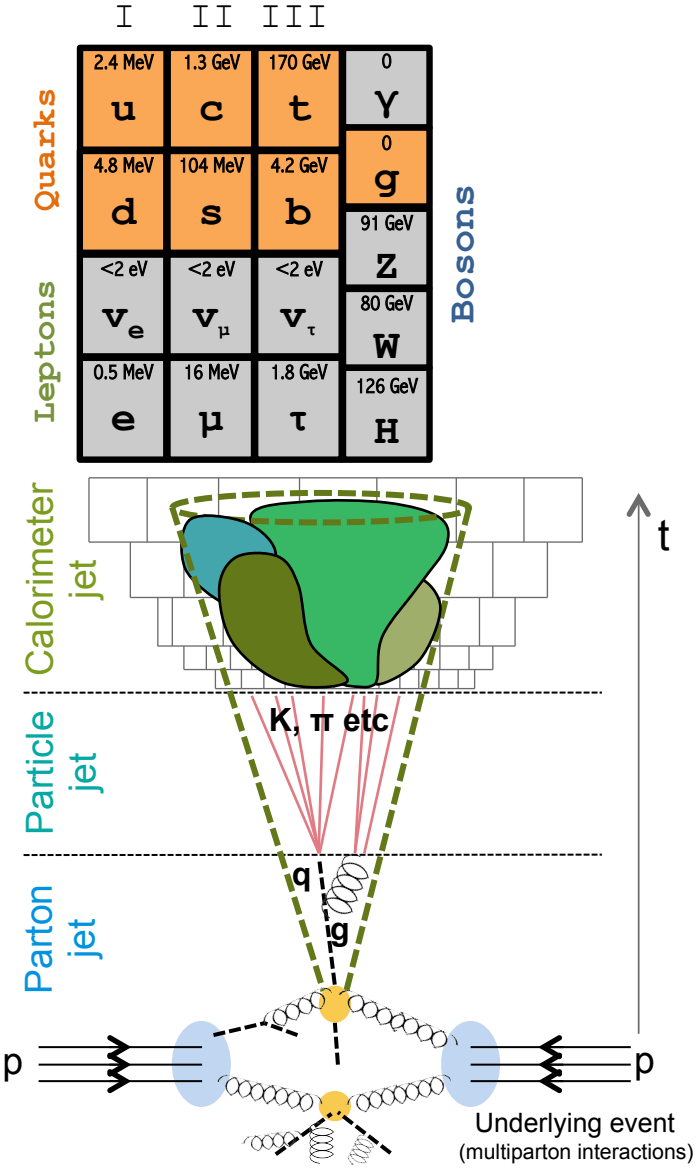
RECONSTRUCTING PARTICLES

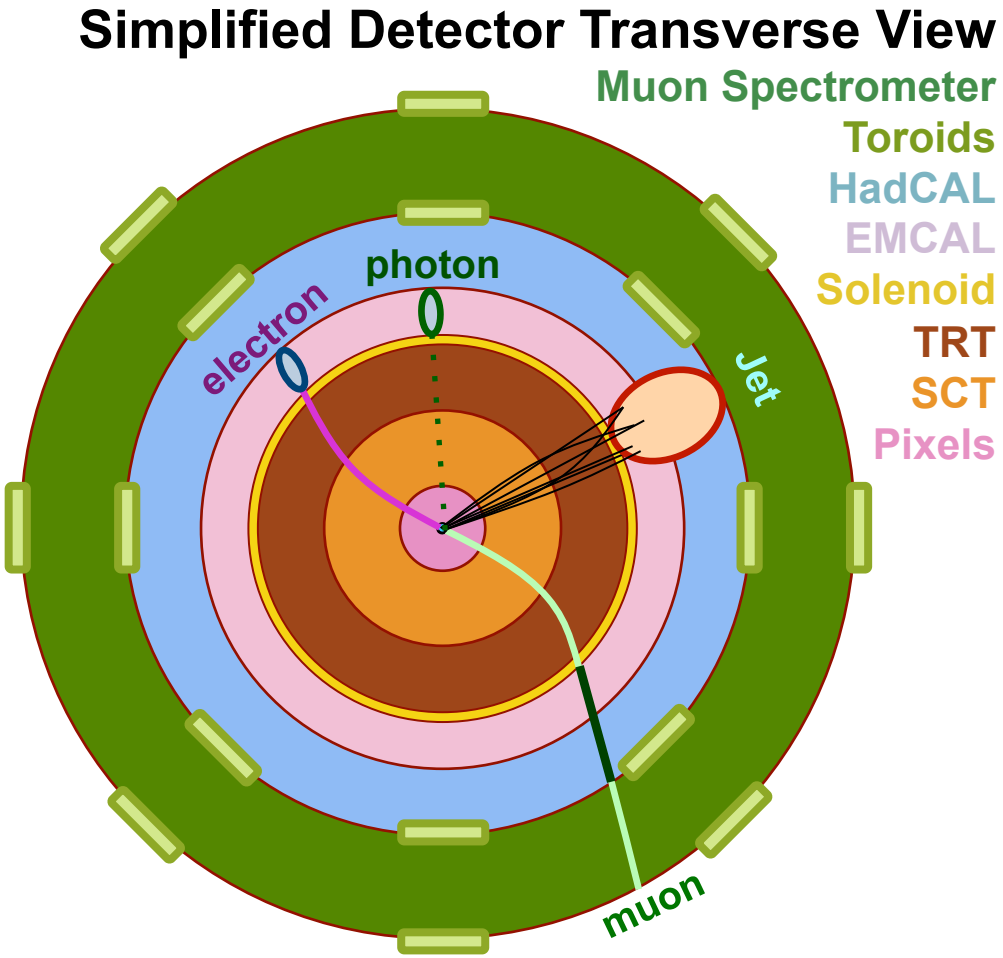
	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2 eV ν_e	<2 eV ν_μ	<2 eV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

Bosons



RECONSTRUCTING PARTICLES



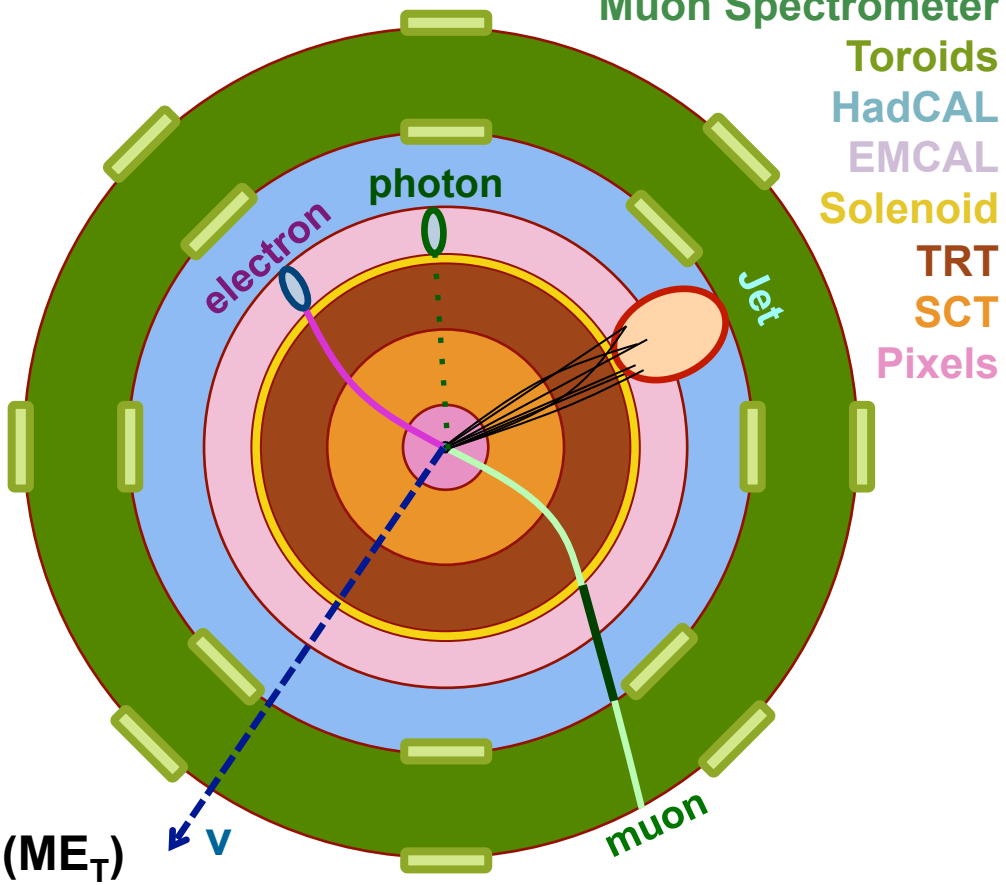


RECONSTRUCTING PARTICLES

	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2 eV ν_e	<2 eV ν_μ	<2 eV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

Bosons

Simplified Detector Transverse View



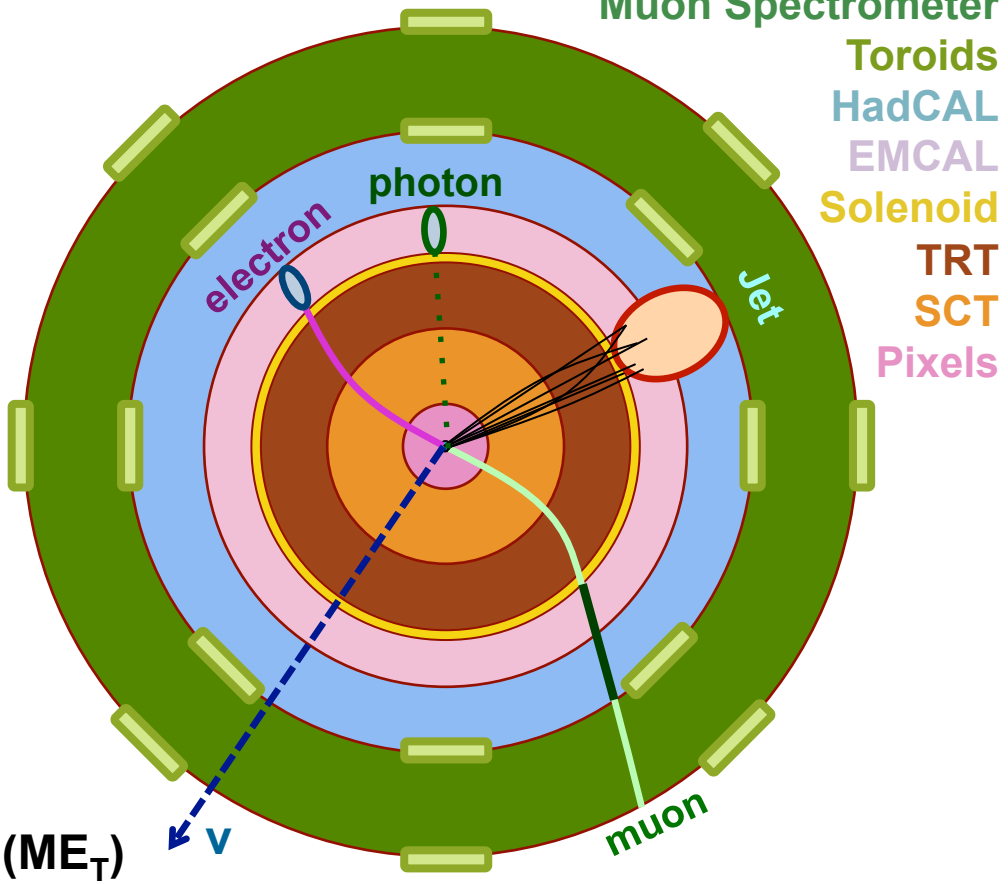
In the transverse plane:

$$\sum \vec{p}_T = 0$$

Missing Transverse Momentum (ME_T)

RECONSTRUCTING PARTICLES

Simplified Detector Transverse View



In the transverse plane:

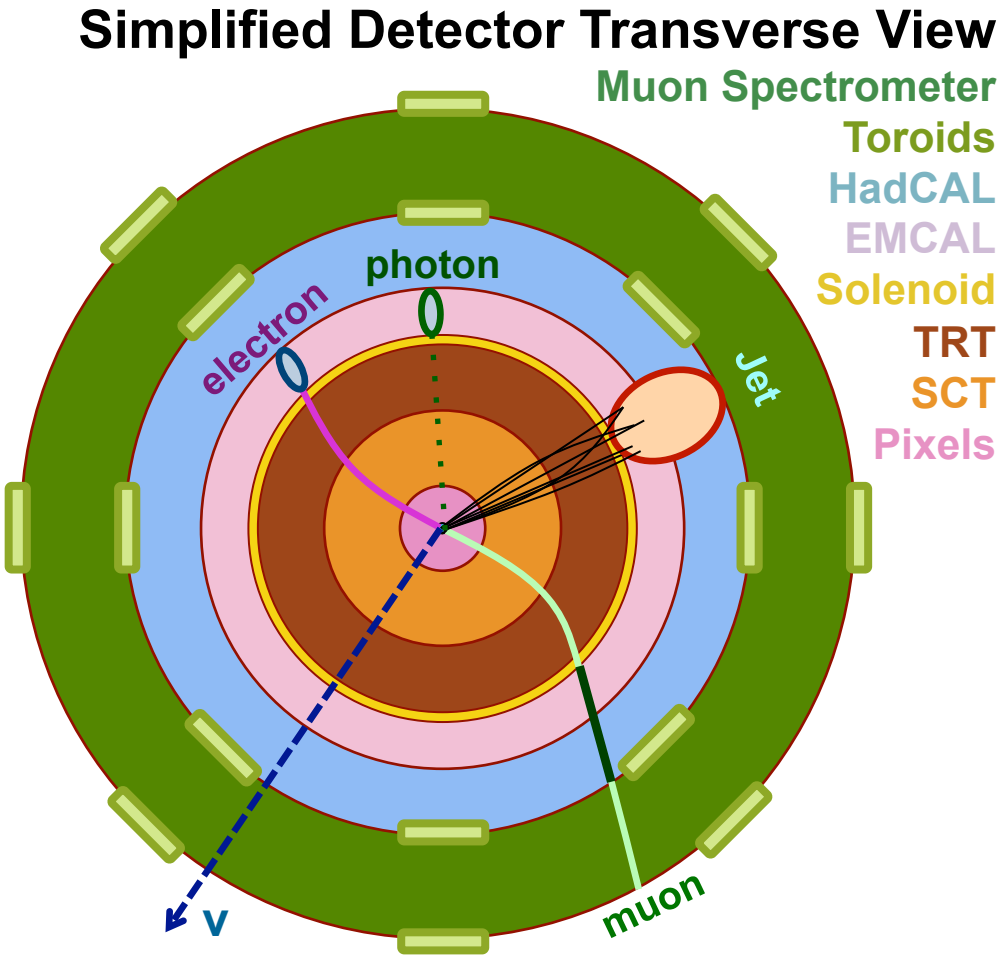
$$\sum \vec{p}_T = 0$$

Missing Transverse Momentum (ME_T)

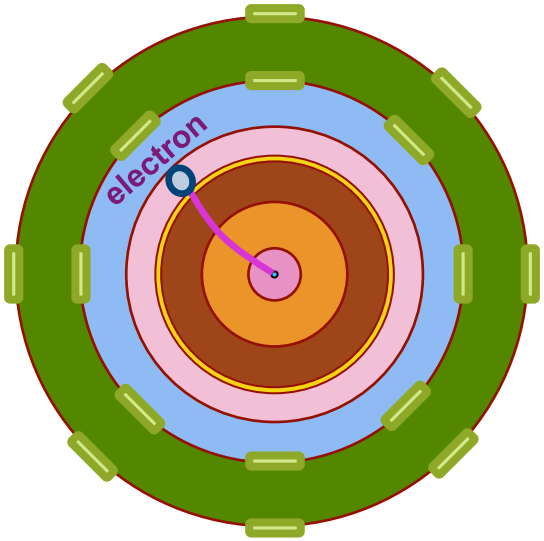
RECONSTRUCTING PARTICLES

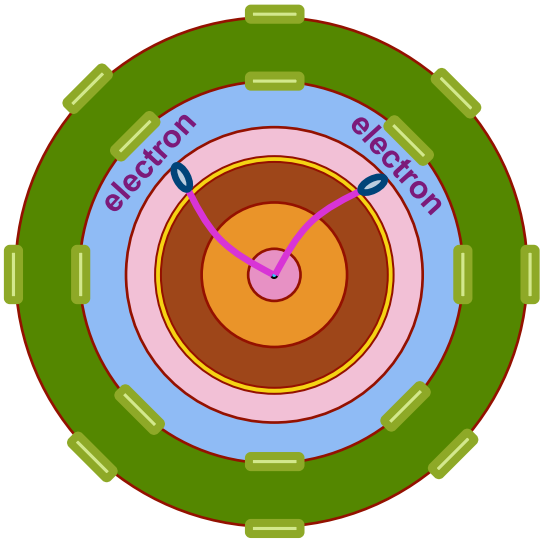
	I	II	III	
Quarks	2.4 MeV u	1.3 GeV c	170 GeV t	0 γ
	4.8 MeV d	104 MeV s	4.2 GeV b	0 g
Leptons	<2 eV ν_e	<2 eV ν_μ	<2 eV ν_τ	91 GeV Z
	0.5 MeV e	16 MeV μ	1.8 GeV τ	80 GeV W
				126 GeV H

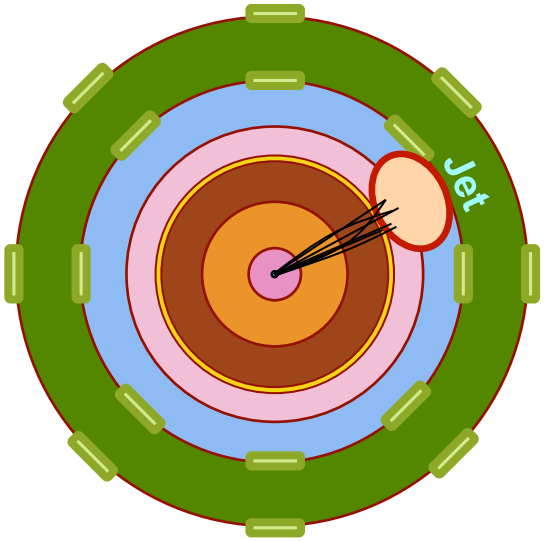
Bosons

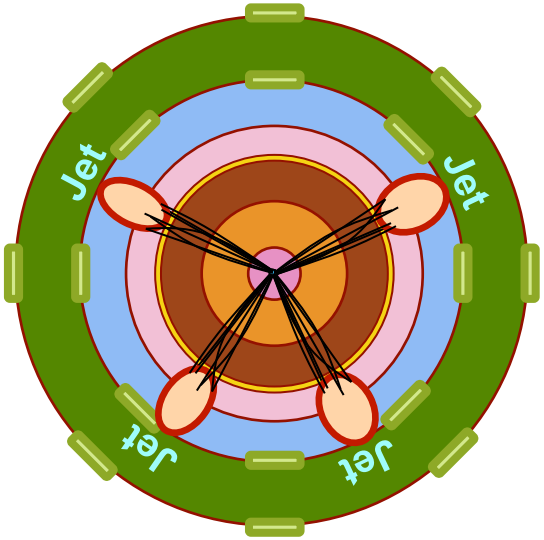


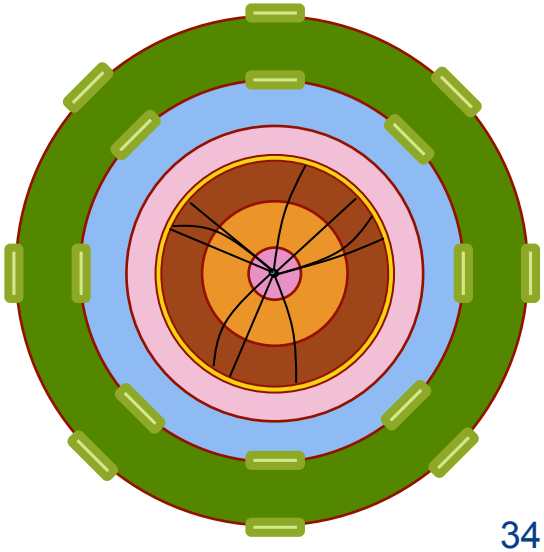
ONLINE RECONSTRUCTION



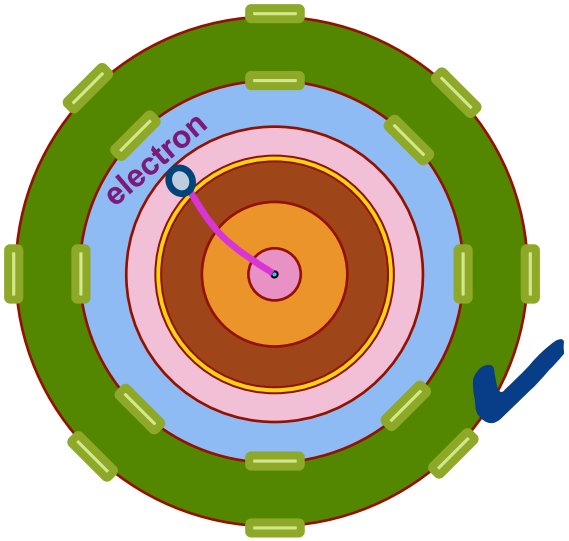


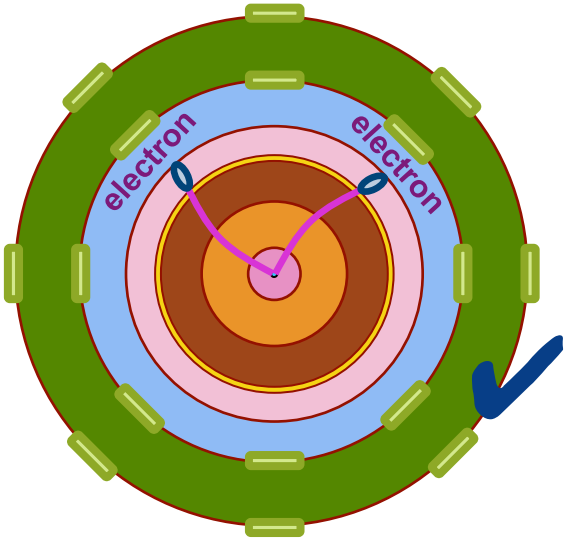


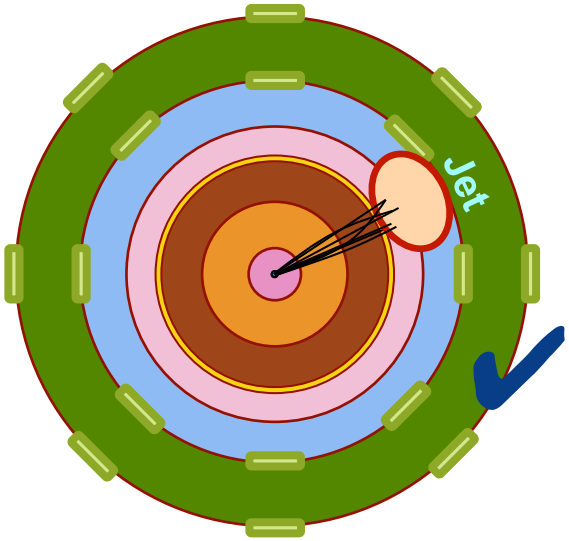


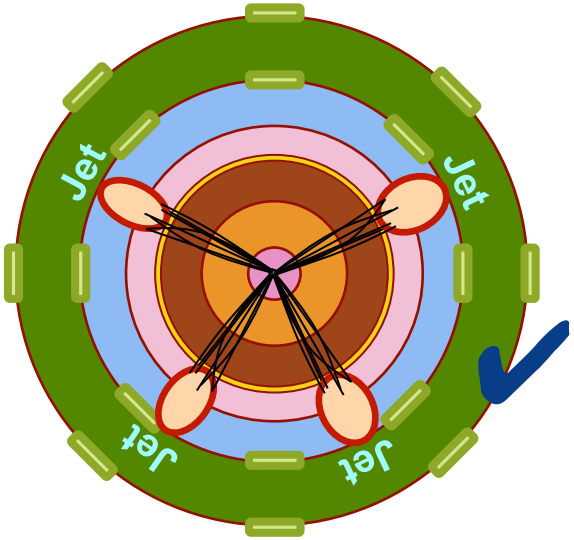


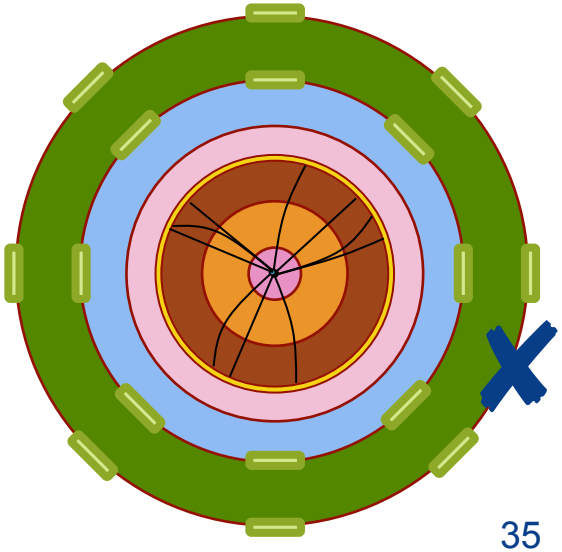
TRIGGERING ON PHYSICS











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

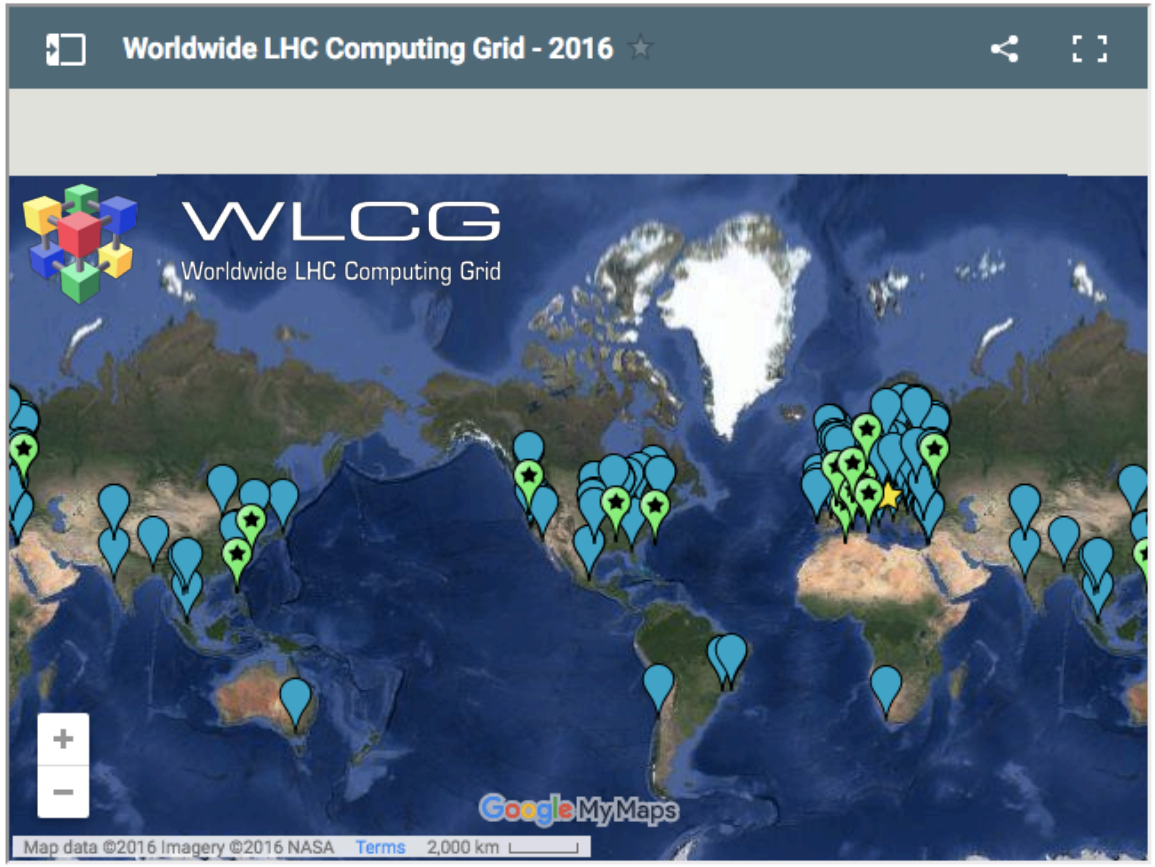
Calibration Streams
events delivering the minimum amount of information for detector calibrations at high rate

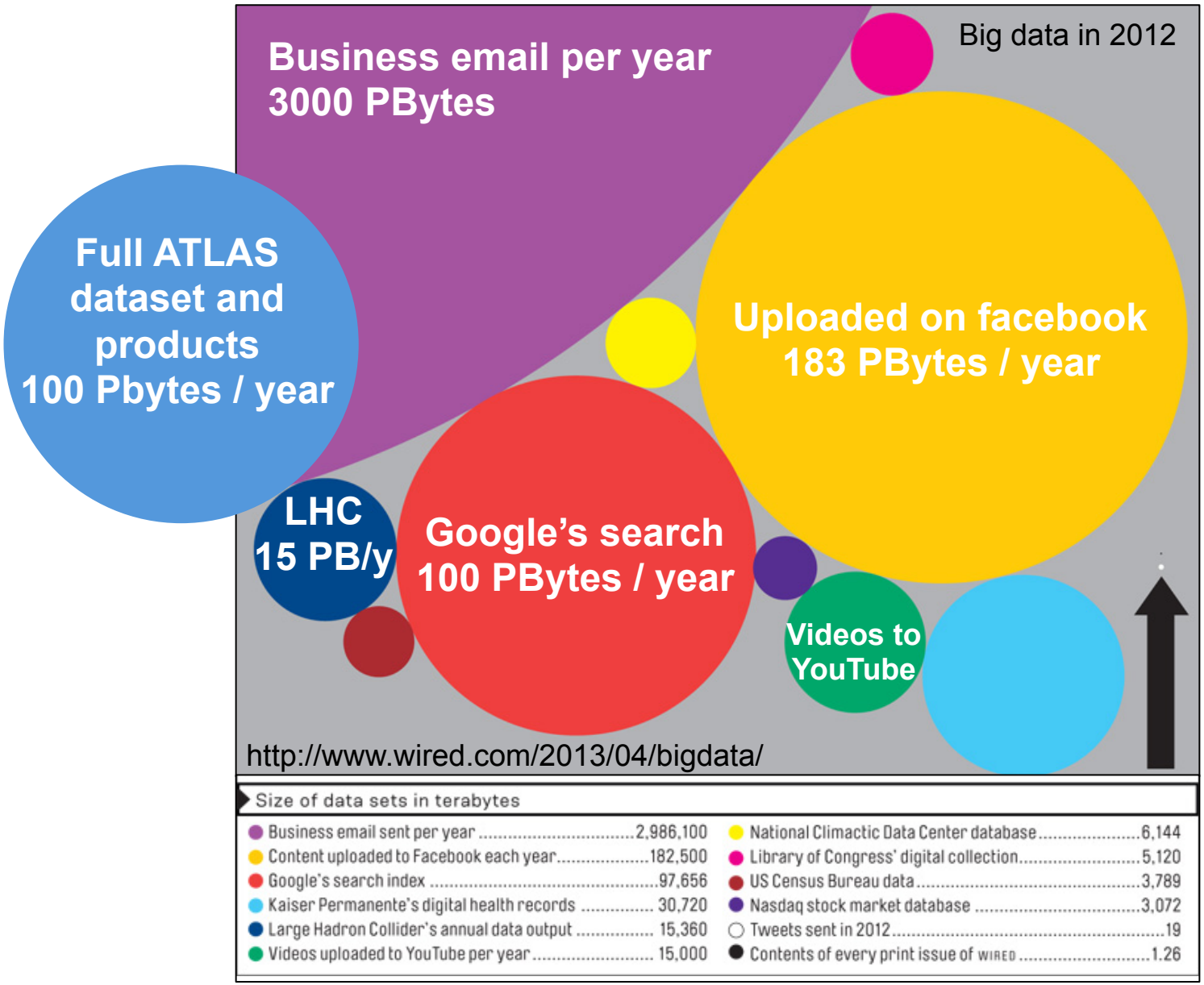
HUGE AMOUNT OF DATA…

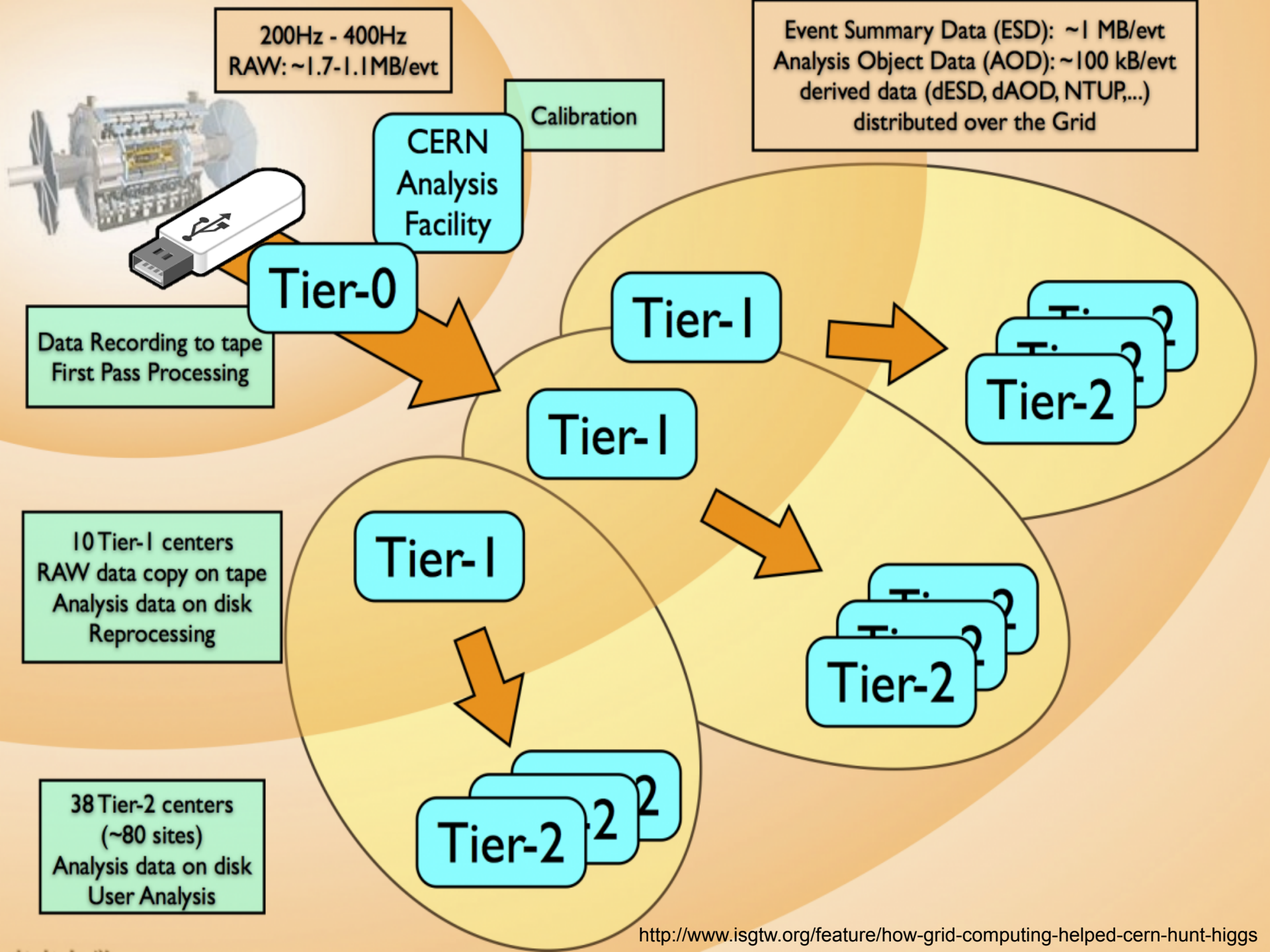
LHC delivered billions of recorded collision events to the LHC experiments from proton-proton and proton-lead collisions in the Run 1 period (2009-2013).

- This translates to ~ 100 PB of data recorded at CERN.

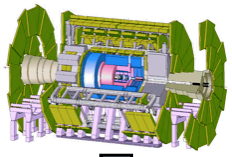
The challenge how to process and analyze the data and produce timely physics results was substantial but in the end resulted in a great success.



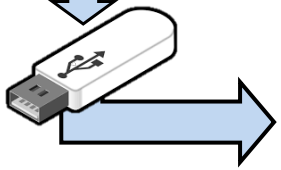

Big data in 2012
Business email per year
3000 PBytes
Full ATLAS
dataset and
products
100 Pbytes / year
Uploaded on facebook
183 PBytes / year
LHC
15 PB/y
Google's search
100 PBytes / year
Videos to
YouTube
http://www.wired.com/2013/04/bigdata/
Size of data sets in terabytes
Business email sent per year 2,986,100
Content uploaded to Facebook each year 182,500
Google's search index 97,656
Kaiser Permanente's digital health records 30,720
Large Hadron Collider's annual data output 15,360
Videos uploaded to YouTube per year 15,000
National Climactic Data Center database 6,144
Library of Congress' digital collection 5,120
US Census Bureau data 3,789
Nasdaq stock market database 3,072
Tweets sent in 2012 19
Contents of every print issue of WIRED 1.26



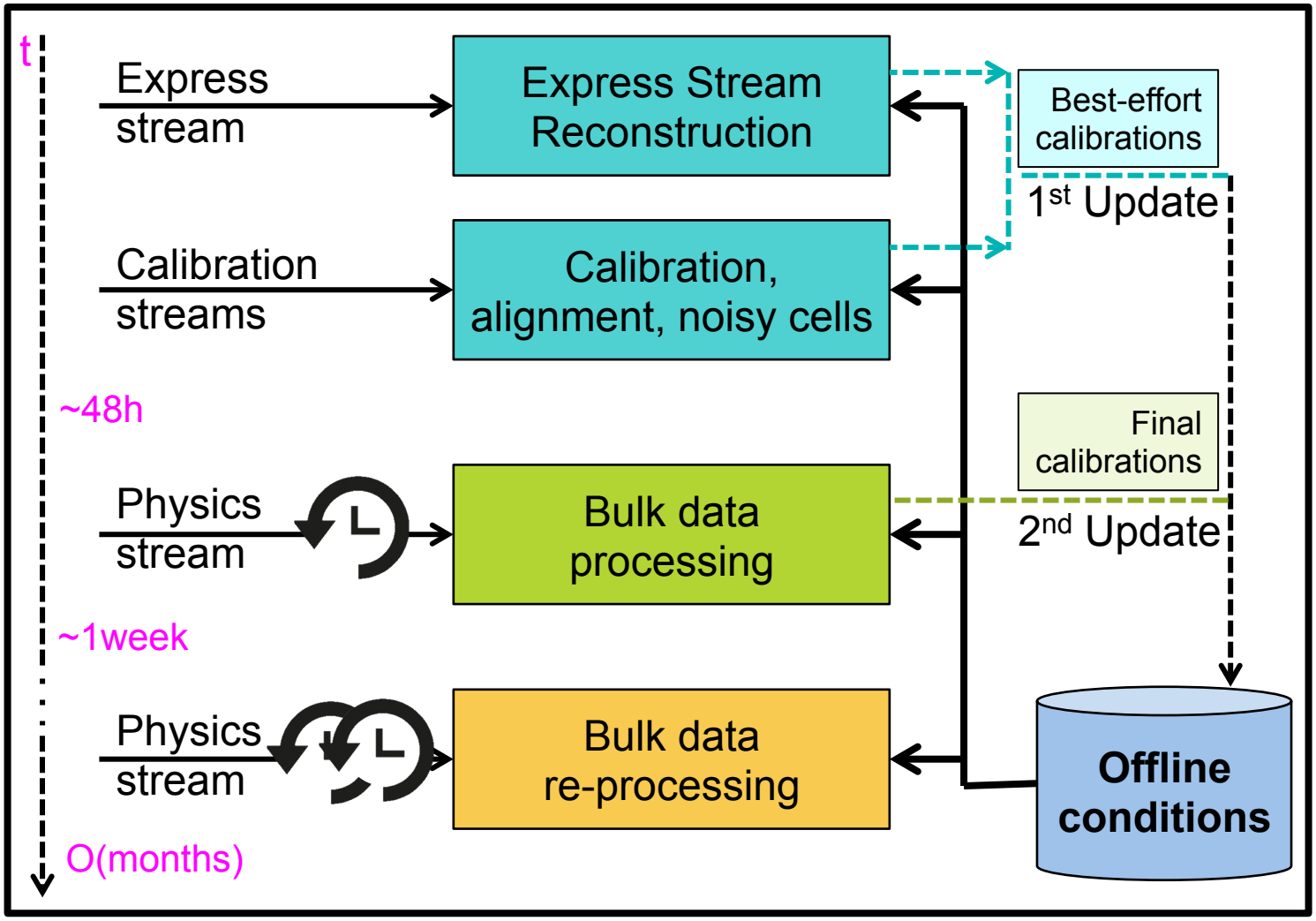
http://www.isgtw.org/feature/how-grid-computing-helped-cern-hunt-higgs



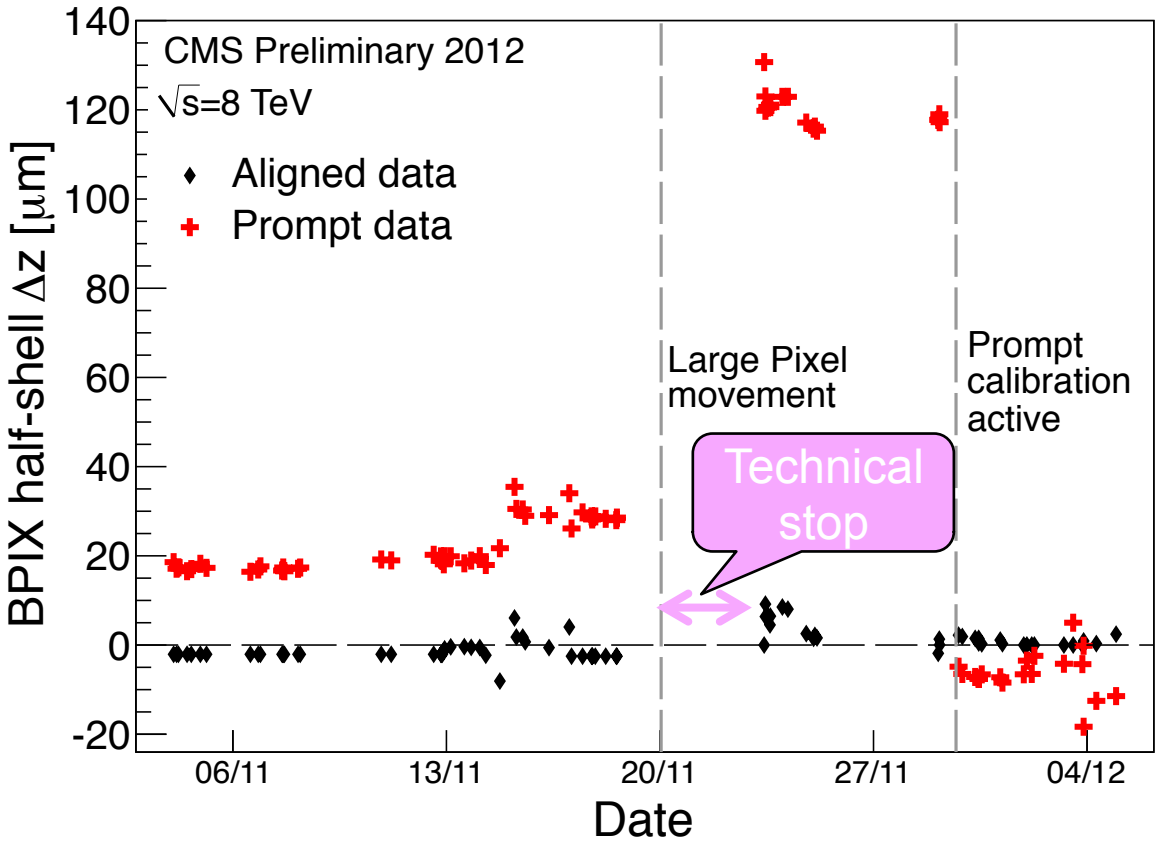
DAQ



THE EVENT AT TIER0

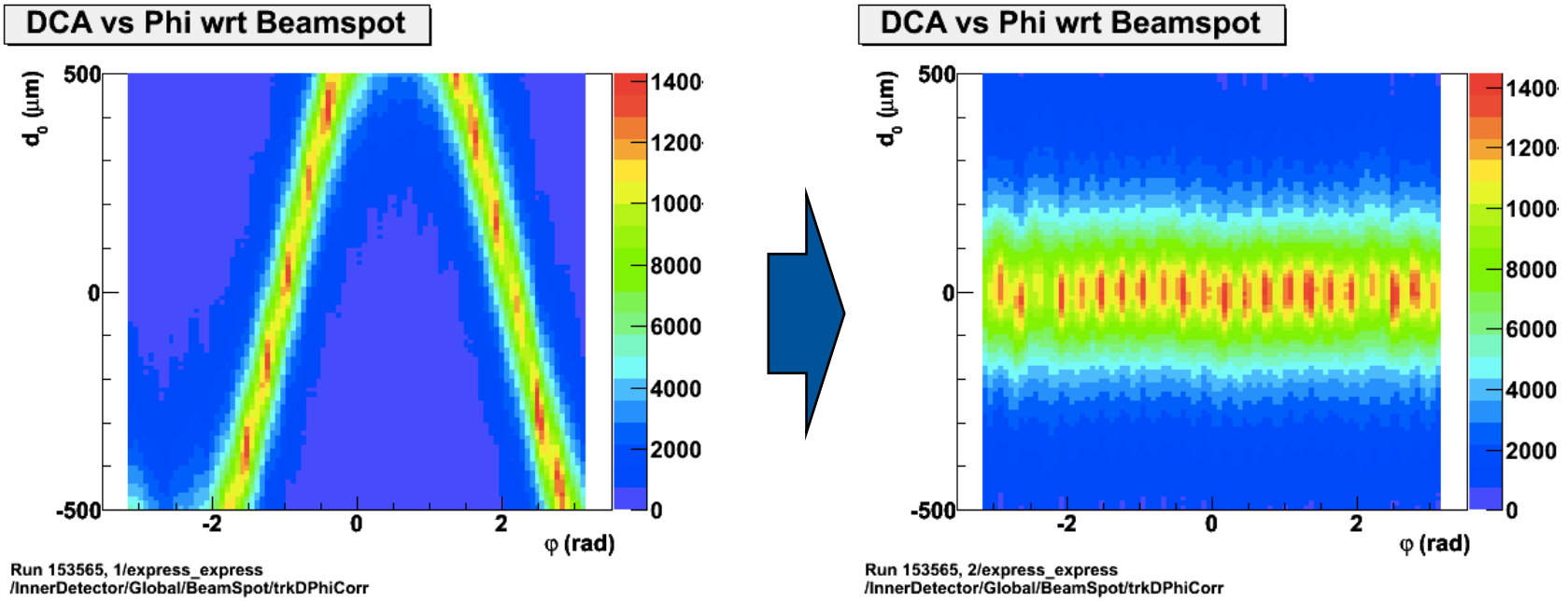


E.G. ALIGNMENT



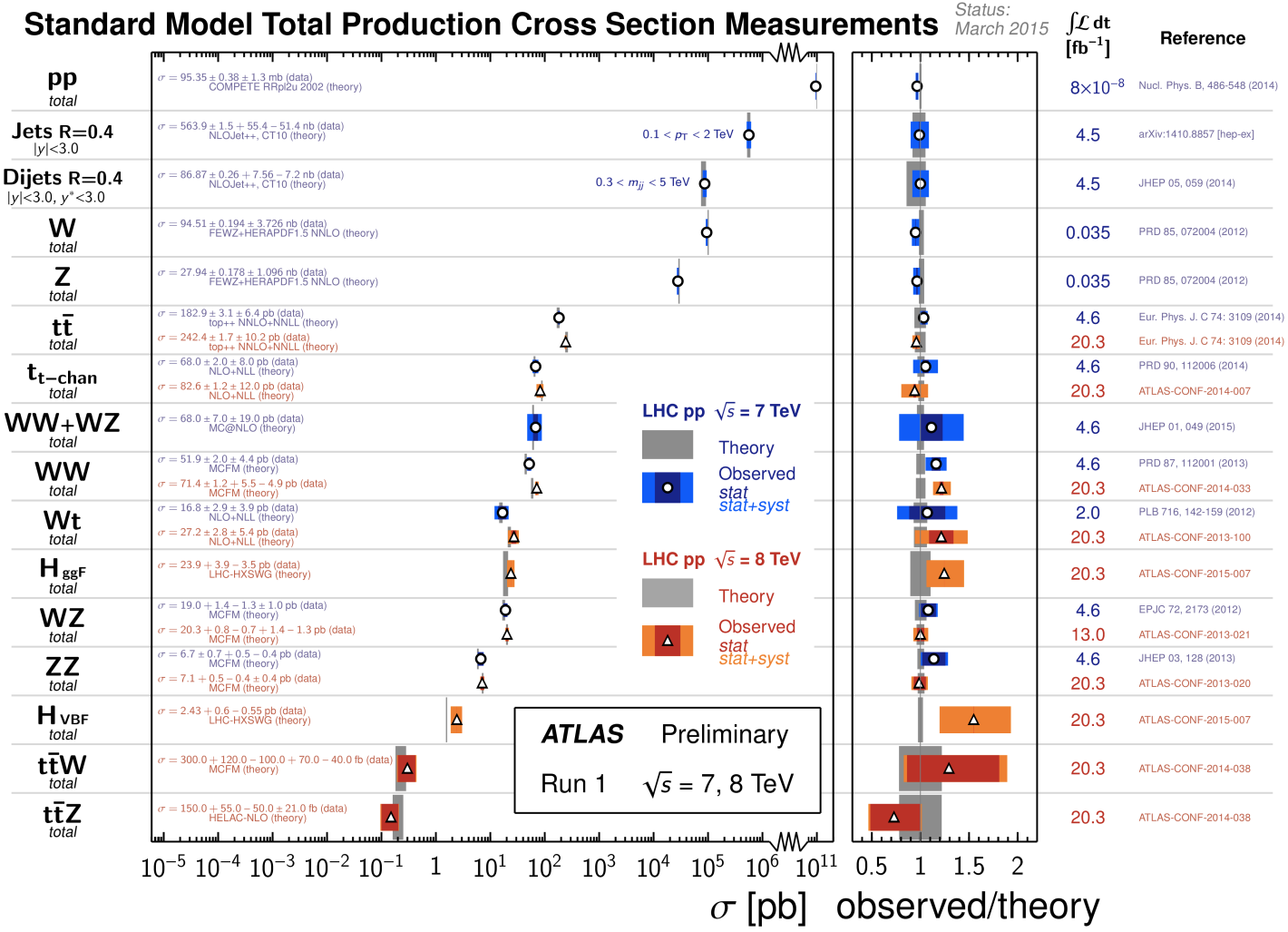
Day-by-day value of the relative longitudinal shift between the two half-shells of the BPIX as measured with the primary vertex residuals, for the last month of pp data taking in 2012.

E.G. BEAMSPOT



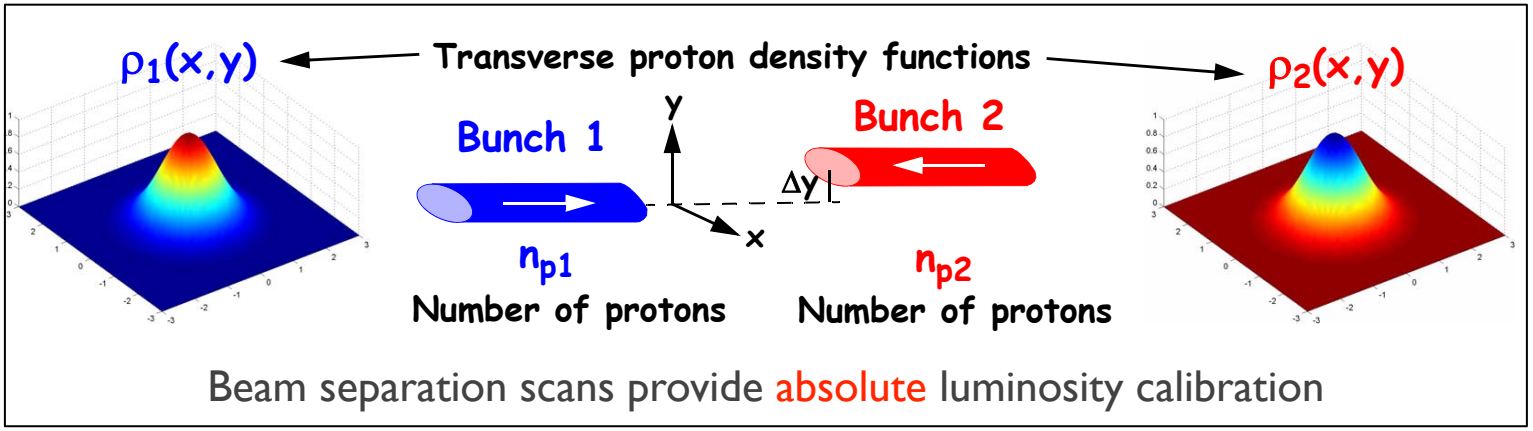
d0 vs phi with respect to the beam spot. For a correctly determined beam spot, this plot should be flat. For the first processing of the **express stream**, the beam spot is not yet known and therefore large variations as in this example are expected. In **bulk reconstruction** this effect is corrected.

“FINAL” CALIBRATION



LUMINOSITY DETERMINATION

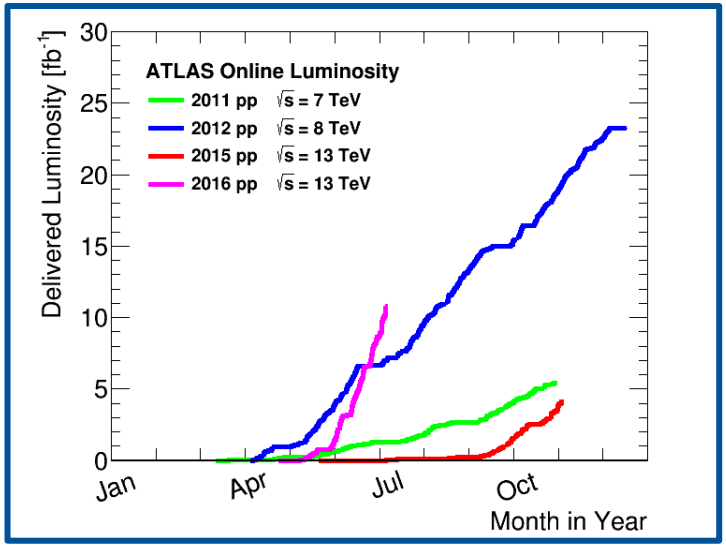
- A measurement of the number of collisions per cm^2 and second.
- Multiple methods used for determining luminosity: reducing uncertainties.
- Normalization is done with beam-separation scan (Van-der-Meer scan). Requires careful control of beam parameters.

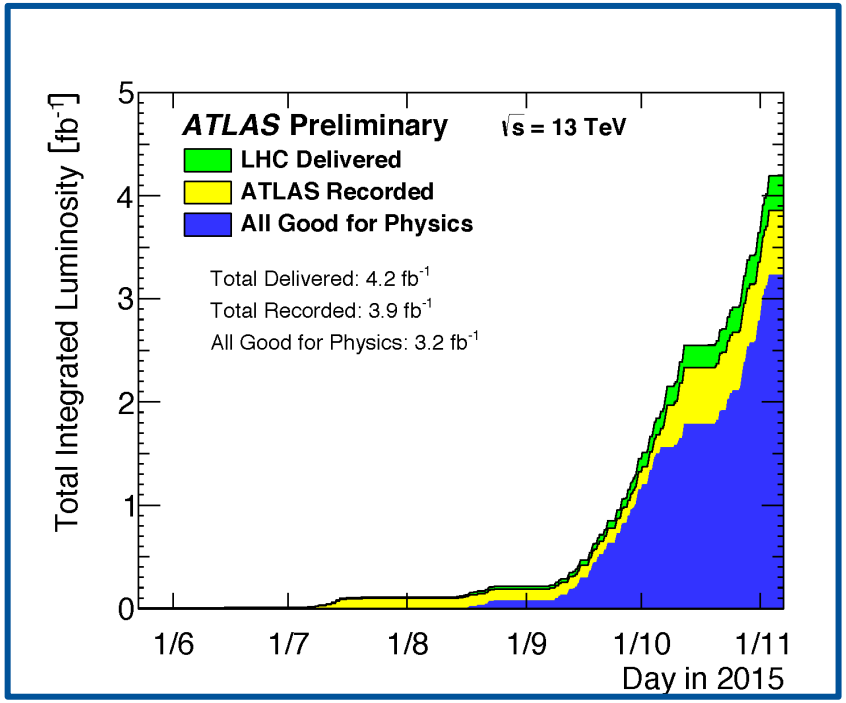


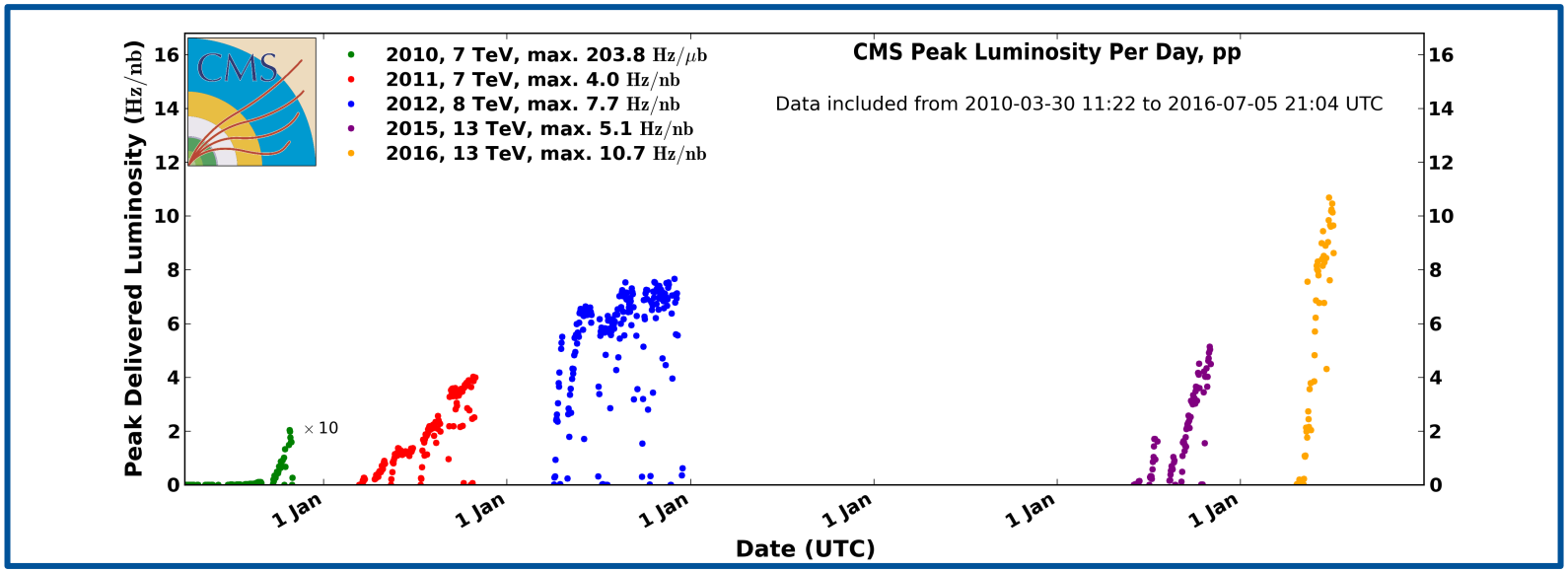
From http://cds.cern.ch/record/1490292/files/ATL-DAPR-SLIDE-2012-627.pdf

- Result: luminosity measurement with very small uncertainties (order of few %) with very fast turn-around time.

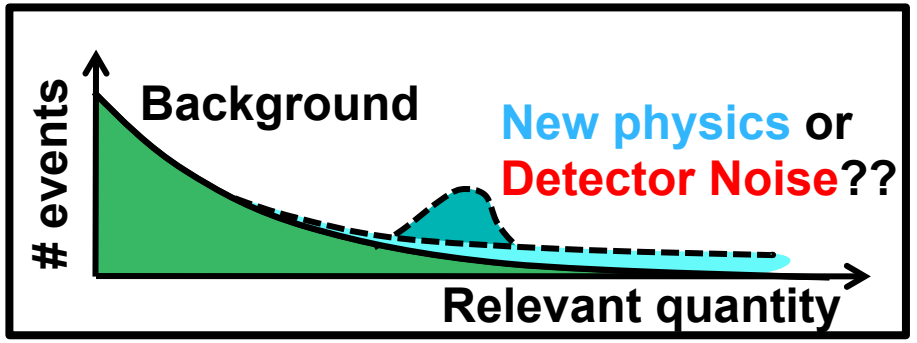
LUMINOSITY


ATLAS Online Luminosity
2011 pp √s = 7 TeV
2012 pp √s = 8 TeV
2015 pp √s = 13 TeV
2016 pp √s = 13 TeV
Delivered Luminosity [fb-1]
Jan
Apr
Jul
Oct
Month in Year


ATLAS Preliminary
√s = 13 TeV
LHC Delivered
ATLAS Recorded
All Good for Physics
Total Delivered: 4.2 fb-1
Total Recorded: 3.9 fb-1
All Good for Physics: 3.2 fb-1
Total Integrated Luminosity [fb-1]
1/6
1/7
1/8
1/9
1/10
1/11
Day in 2015


CMS
2010, 7 TeV, max. 203.8 Hz/μb
2011, 7 TeV, max. 4.0 Hz/nb
2012, 8 TeV, max. 7.7 Hz/nb
2015, 13 TeV, max. 5.1 Hz/nb
2016, 13 TeV, max. 10.7 Hz/nb
CMS Peak Luminosity Per Day, pp
Data included from 2010-03-30 11:22 to 2016-07-05 21:04 UTC
Peak Delivered Luminosity (Hz/nb)
× 10
1 Jan
Date (UTC)

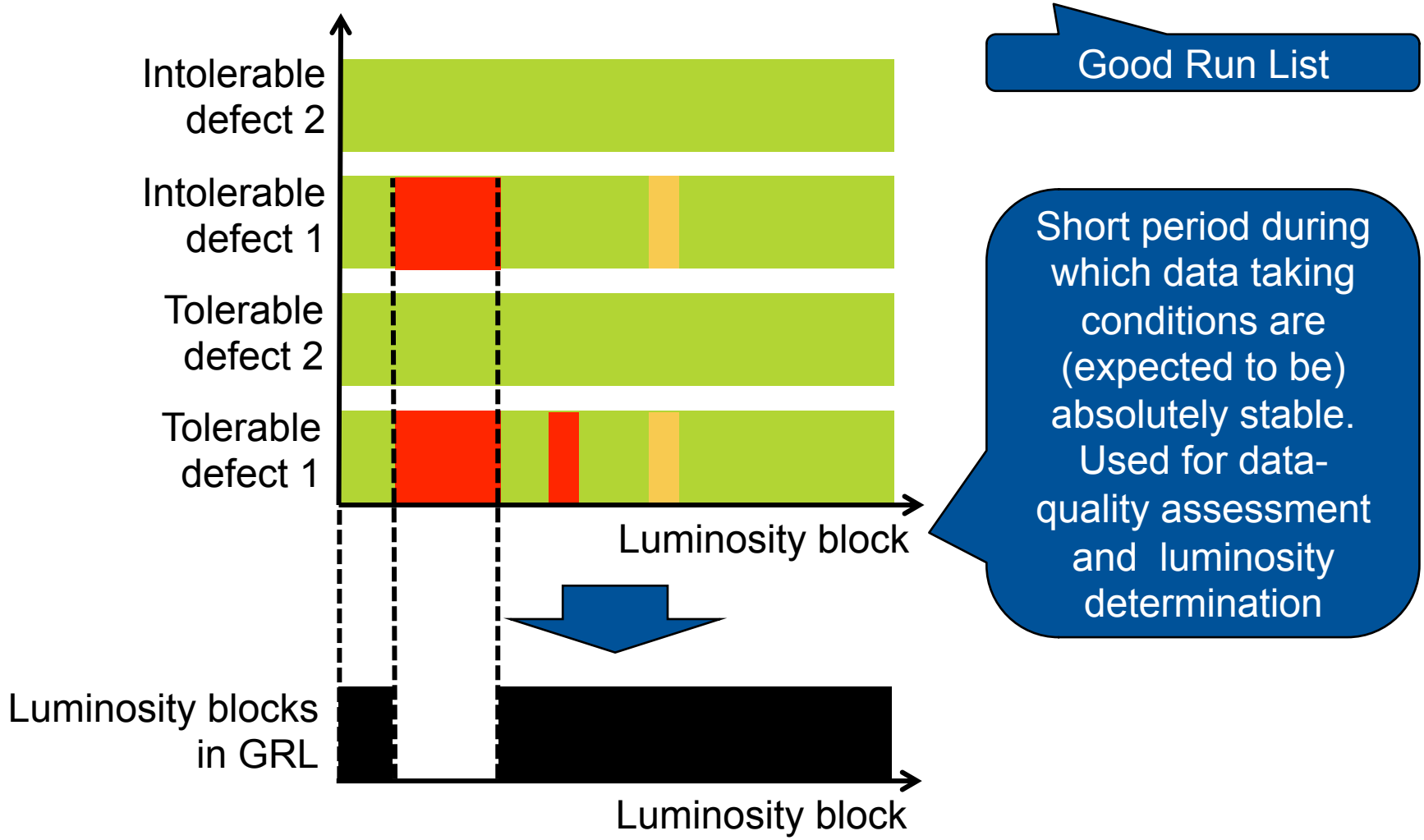
DATA QUALITY



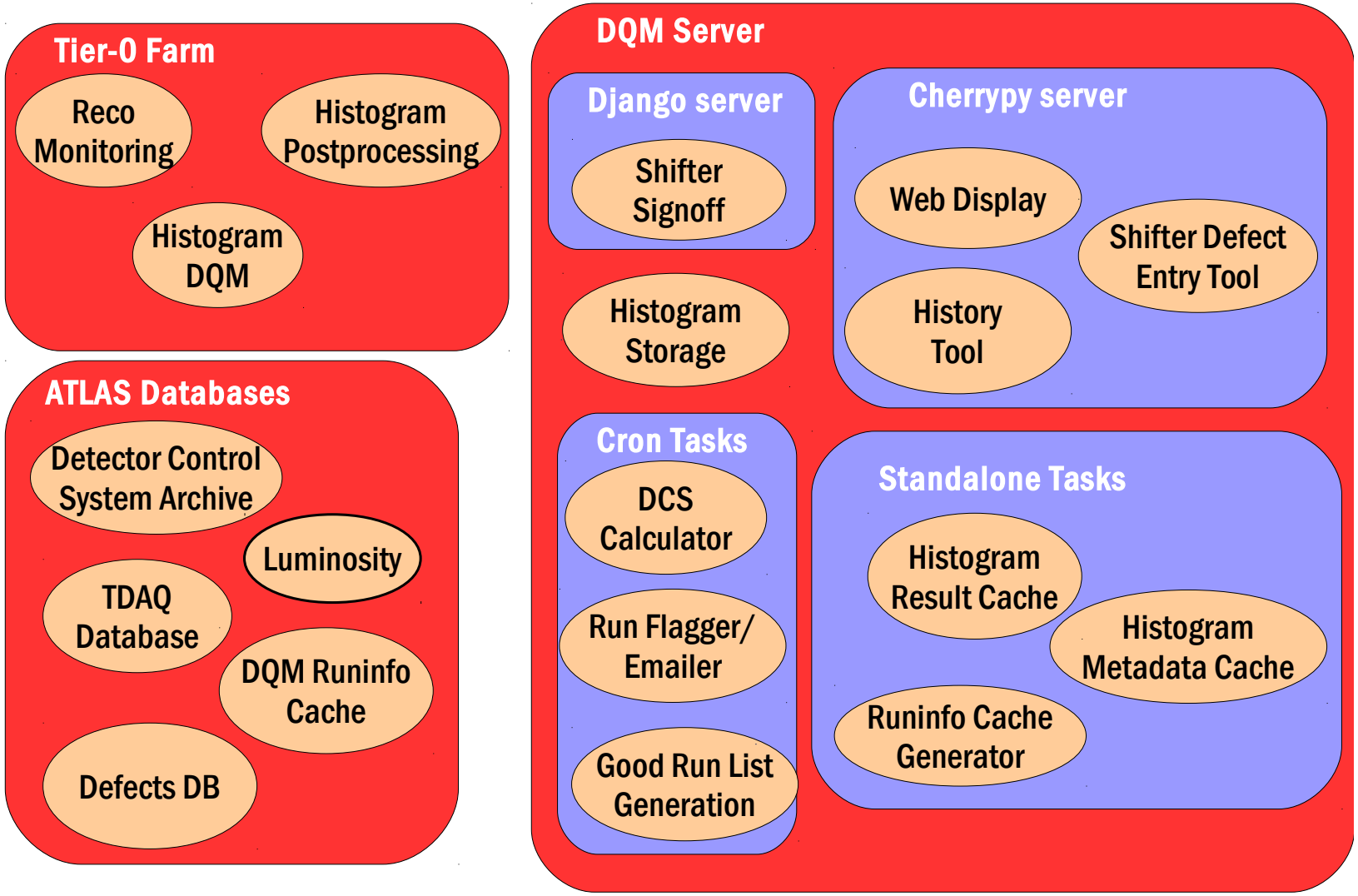
The data we analyze has to follow norms of quality such that our results are trustable.

- **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 AND "GRL"

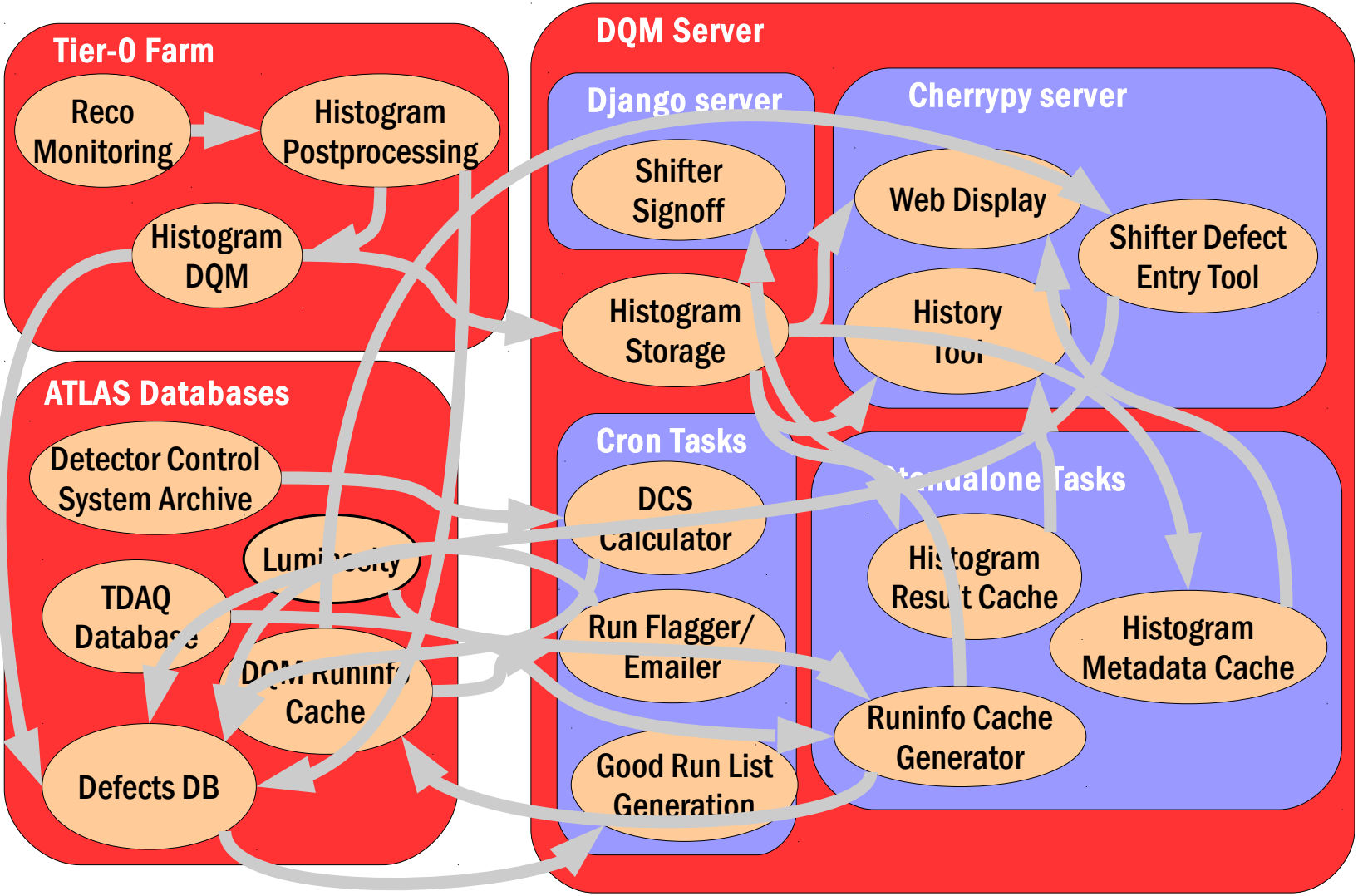


DATA QUALITY



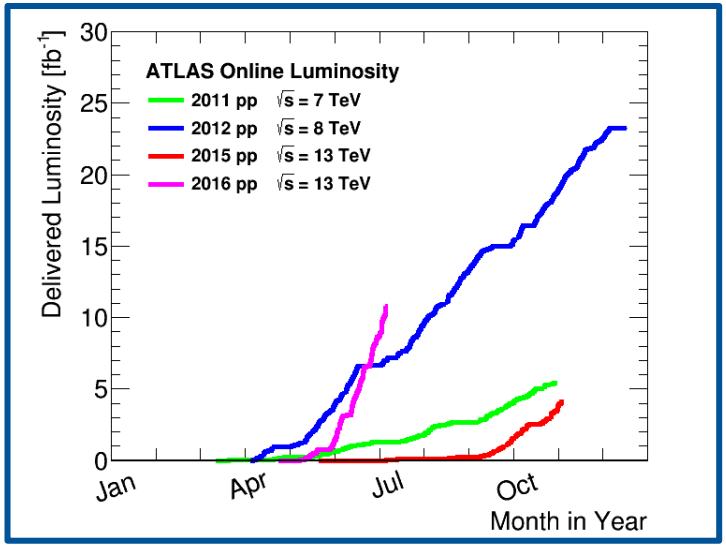
https://cds.cern.ch/record/2008725/files/ATL-SOFT-SLIDE-2015-179.pdf

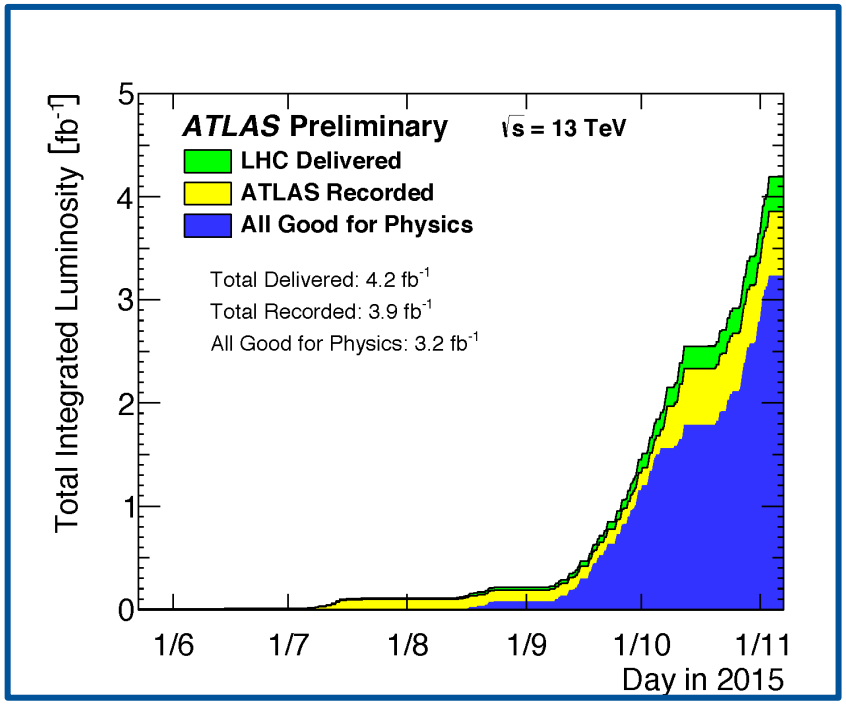
DATA QUALITY

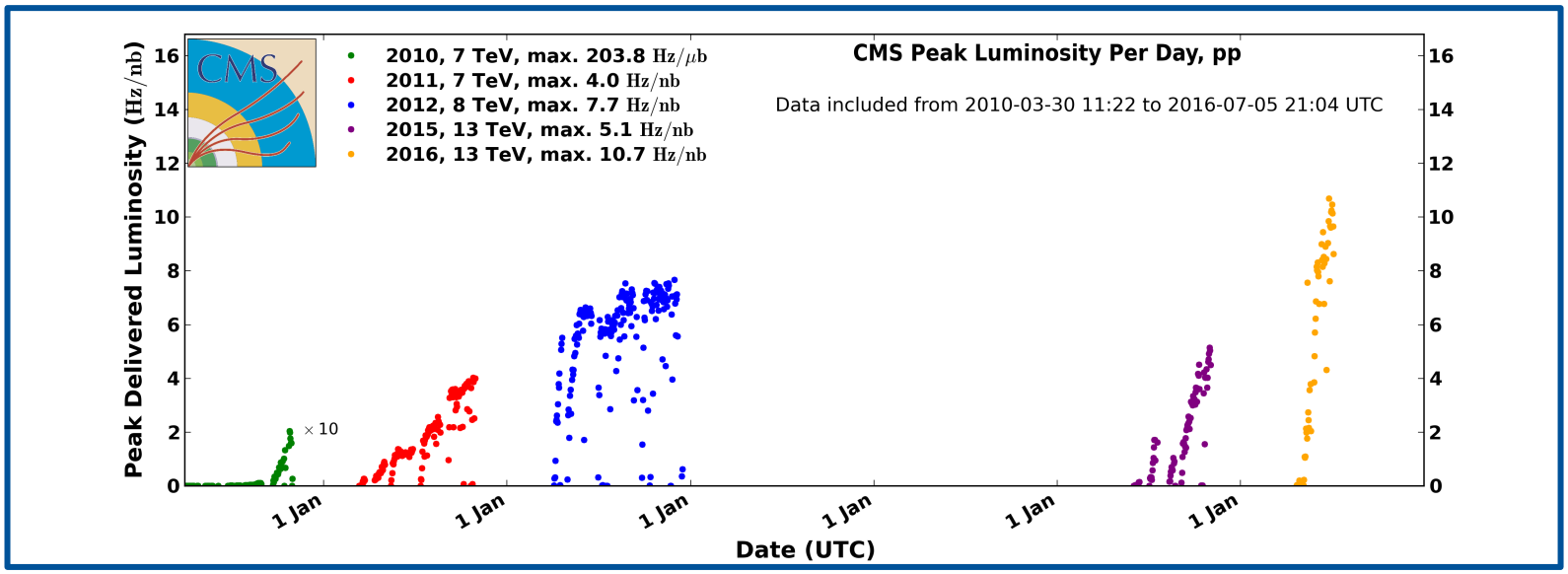


https://cds.cern.ch/record/2008725/files/ATL-SOFT-SLIDE-2015-179.pdf

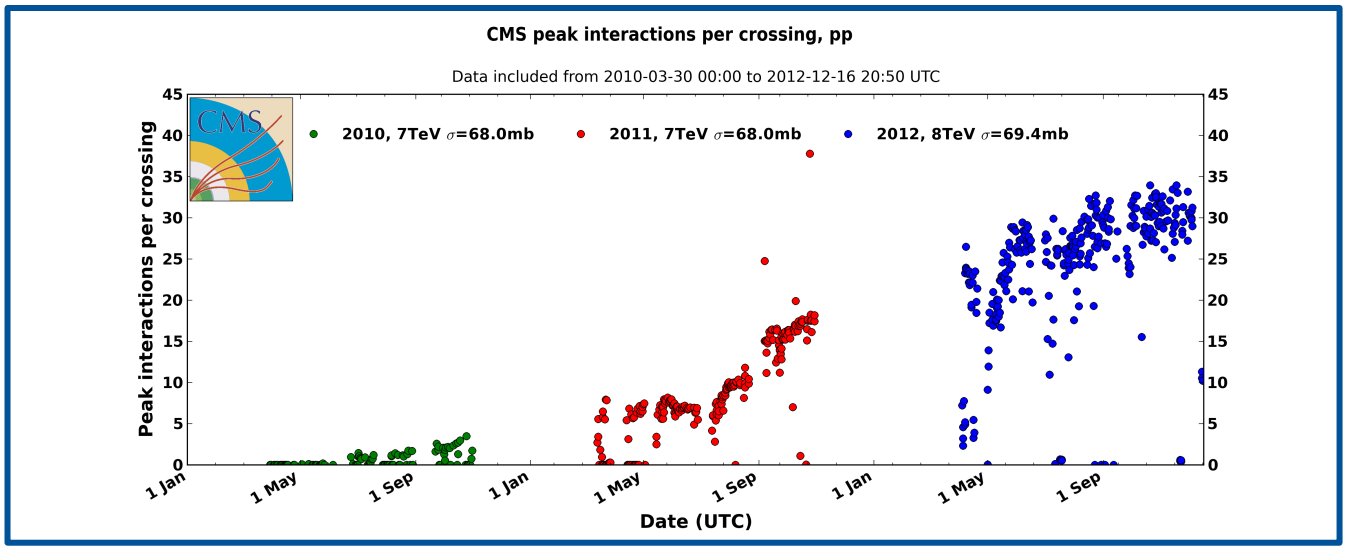
LUMINOSITY

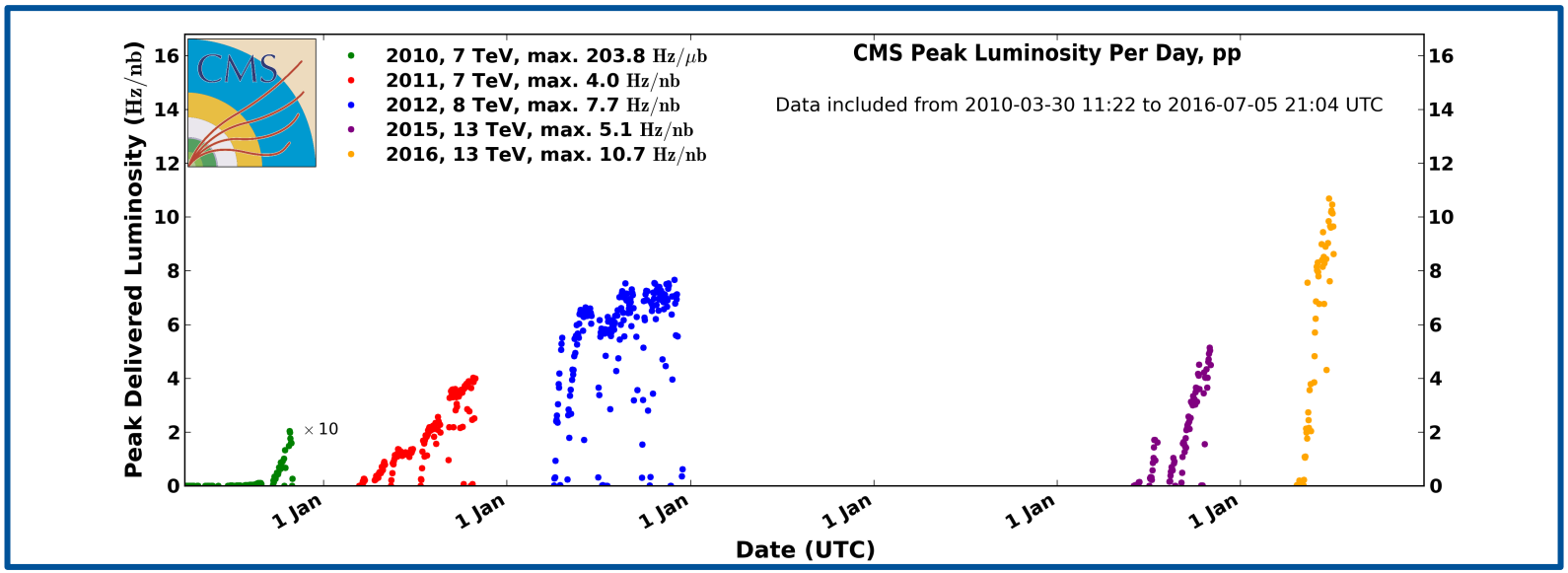

ATLAS Online Luminosity
2011 pp √s = 7 TeV
2012 pp √s = 8 TeV
2015 pp √s = 13 TeV
2016 pp √s = 13 TeV
Delivered Luminosity [fb-1]
Jan
Apr
Jul
Oct
Month in Year


ATLAS Preliminary
√s = 13 TeV
LHC Delivered
ATLAS Recorded
All Good for Physics
Total Delivered: 4.2 fb-1
Total Recorded: 3.9 fb-1
All Good for Physics: 3.2 fb-1
Total Integrated Luminosity [fb-1]
1/6
1/7
1/8
1/9
1/10
1/11
Day in 2015

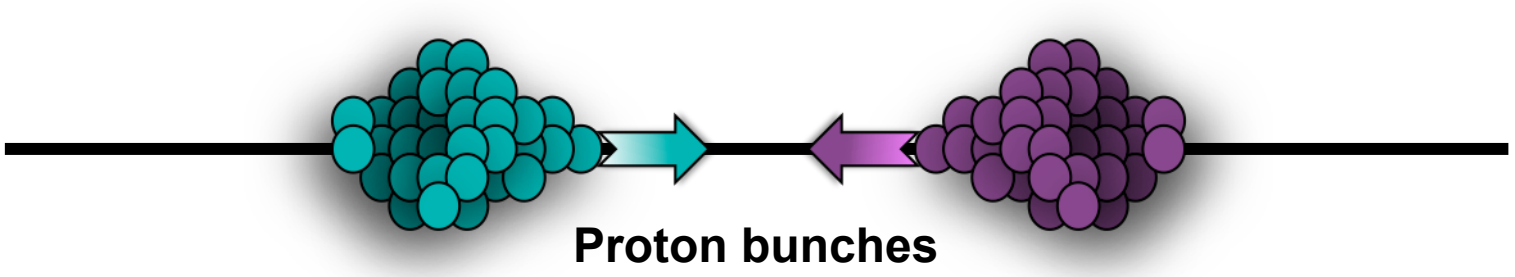

CMS
2010, 7 TeV, max. 203.8 Hz/μb
2011, 7 TeV, max. 4.0 Hz/nb
2012, 8 TeV, max. 7.7 Hz/nb
2015, 13 TeV, max. 5.1 Hz/nb
2016, 13 TeV, max. 10.7 Hz/nb
CMS Peak Luminosity Per Day, pp
Data included from 2010-03-30 11:22 to 2016-07-05 21:04 UTC
Peak Delivered Luminosity (Hz/nb)
× 10
1 Jan
Date (UTC)

LUMINOSITY

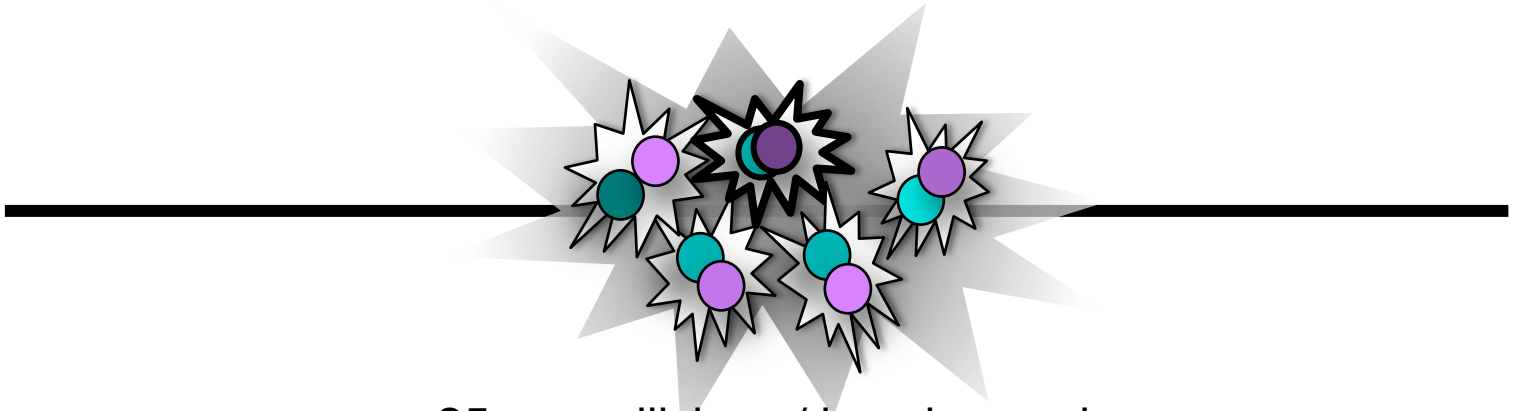

CMS peak interactions per crossing, pp
Data included from 2010-03-30 00:00 to 2012-12-16 20:50 UTC
CMS
2010, 7TeV σ=68.0mb
2011, 7TeV σ=68.0mb
2012, 8TeV σ=69.4mb
Peak interactions per crossing
Date (UTC)
1 Jan
1 May
1 Sep


CMS Peak Luminosity Per Day, pp
Data included from 2010-03-30 11:22 to 2016-07-05 21:04 UTC
CMS
2010, 7 TeV, max. 203.8 Hz/μb
2011, 7 TeV, max. 4.0 Hz/nb
2012, 8 TeV, max. 7.7 Hz/nb
2015, 13 TeV, max. 5.1 Hz/nb
2016, 13 TeV, max. 10.7 Hz/nb
× 10
Peak Delivered Luminosity (Hz/nb)
Date (UTC)
1 Jan

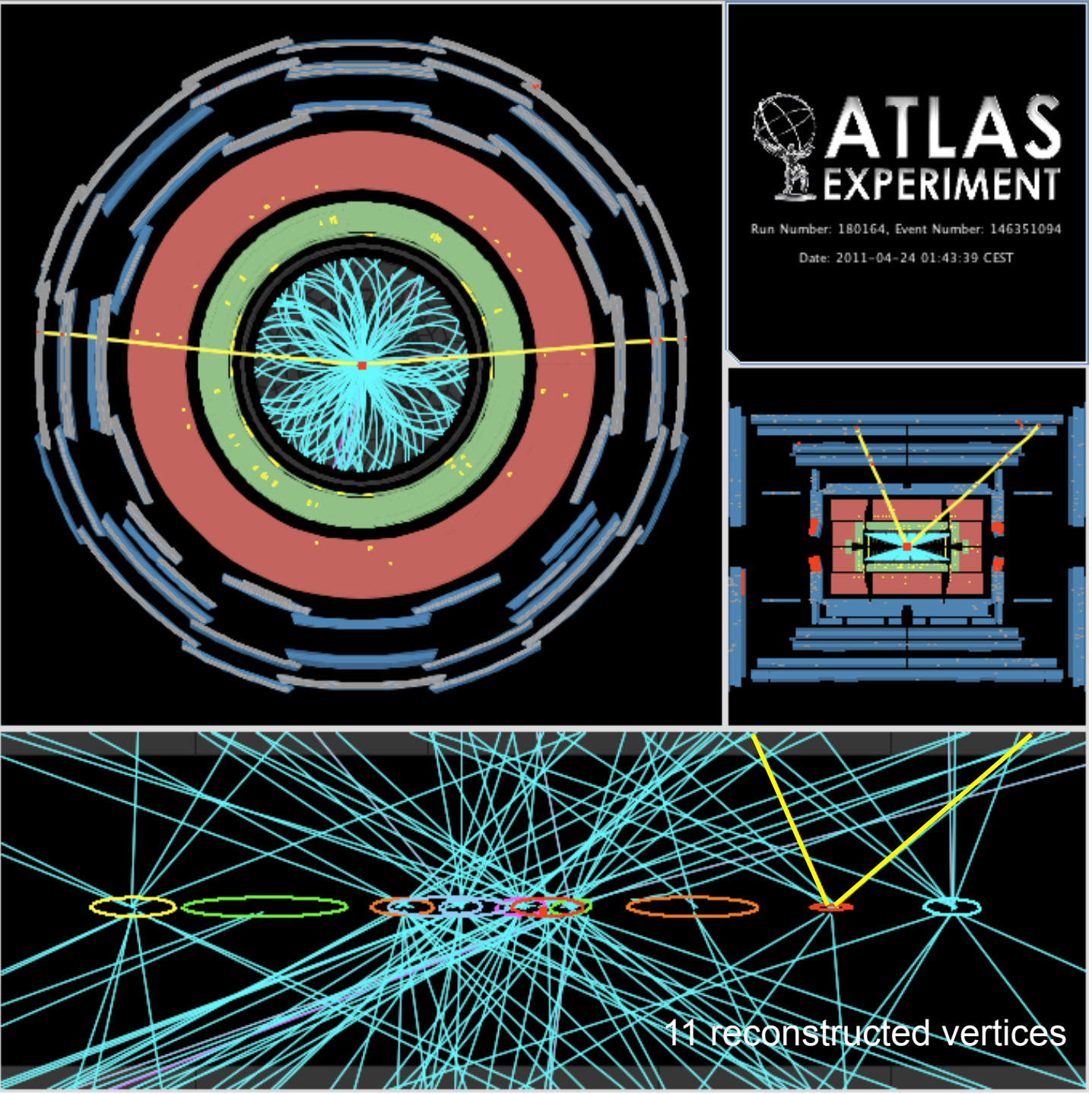
PILE-UP



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

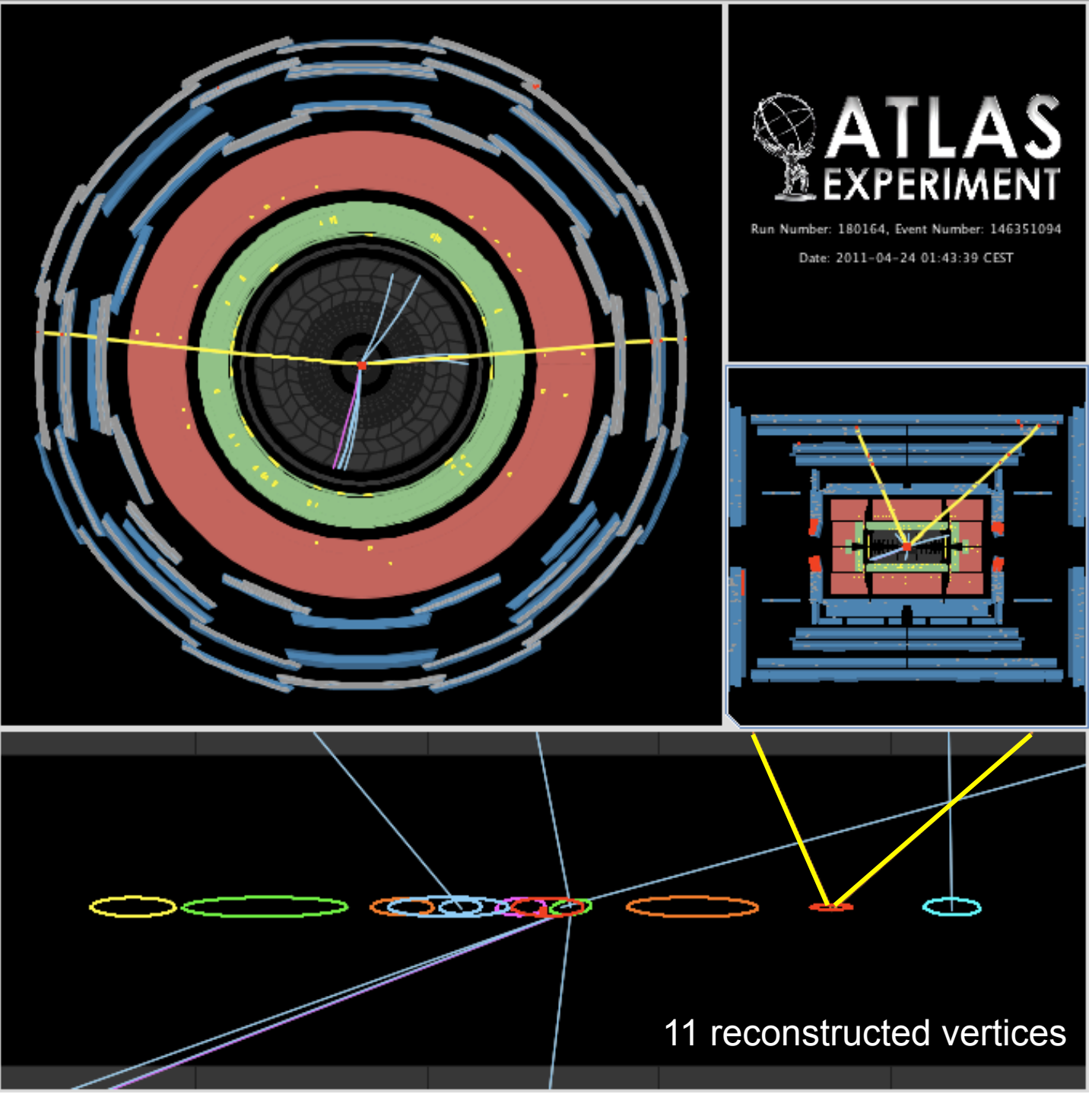


~25 p-p collisions / bunch crossing



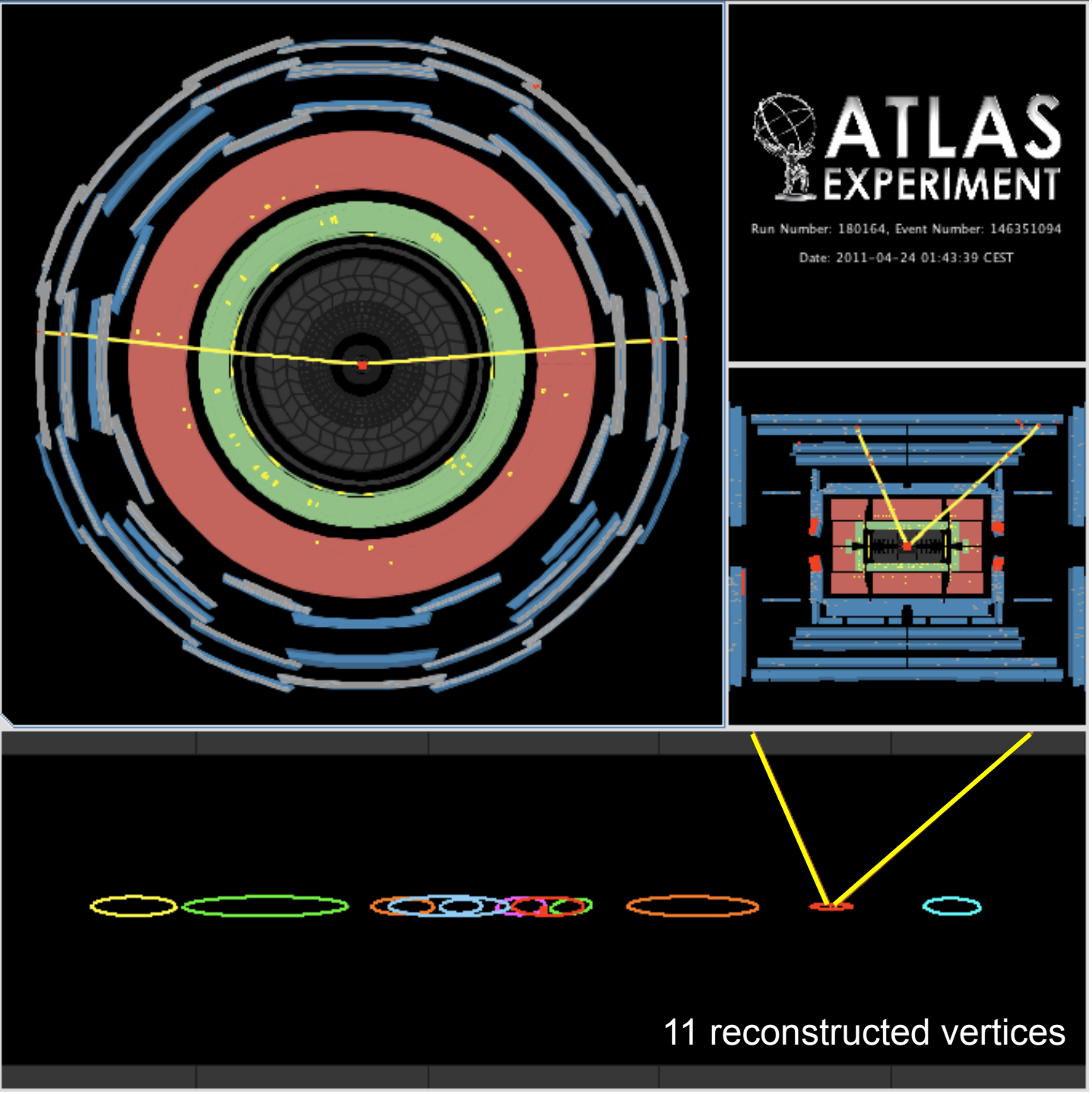
Z->μμ event;
2011 data.

Track pT > 0.5 GeV



Z->μμ event;
2011 data.

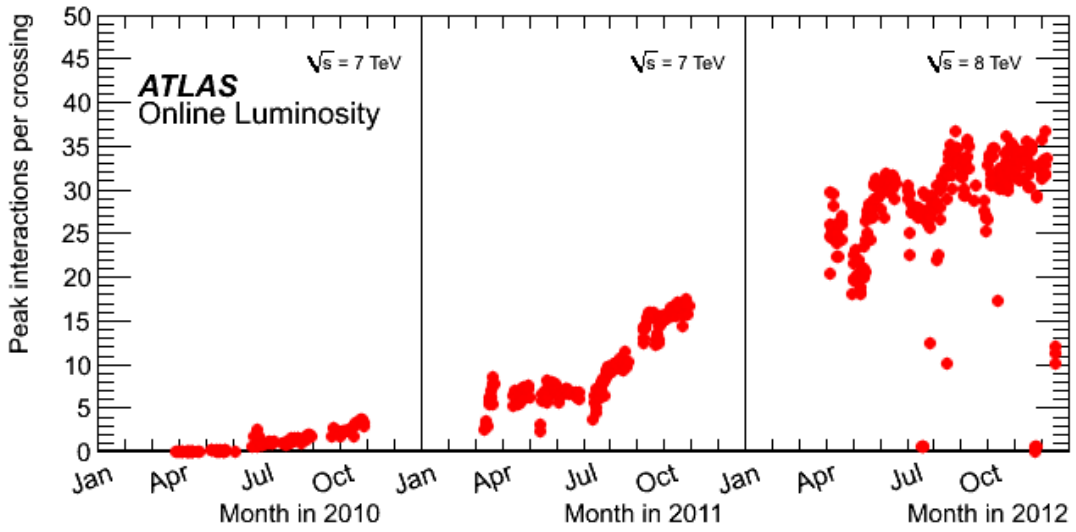
Track pT > 2 GeV

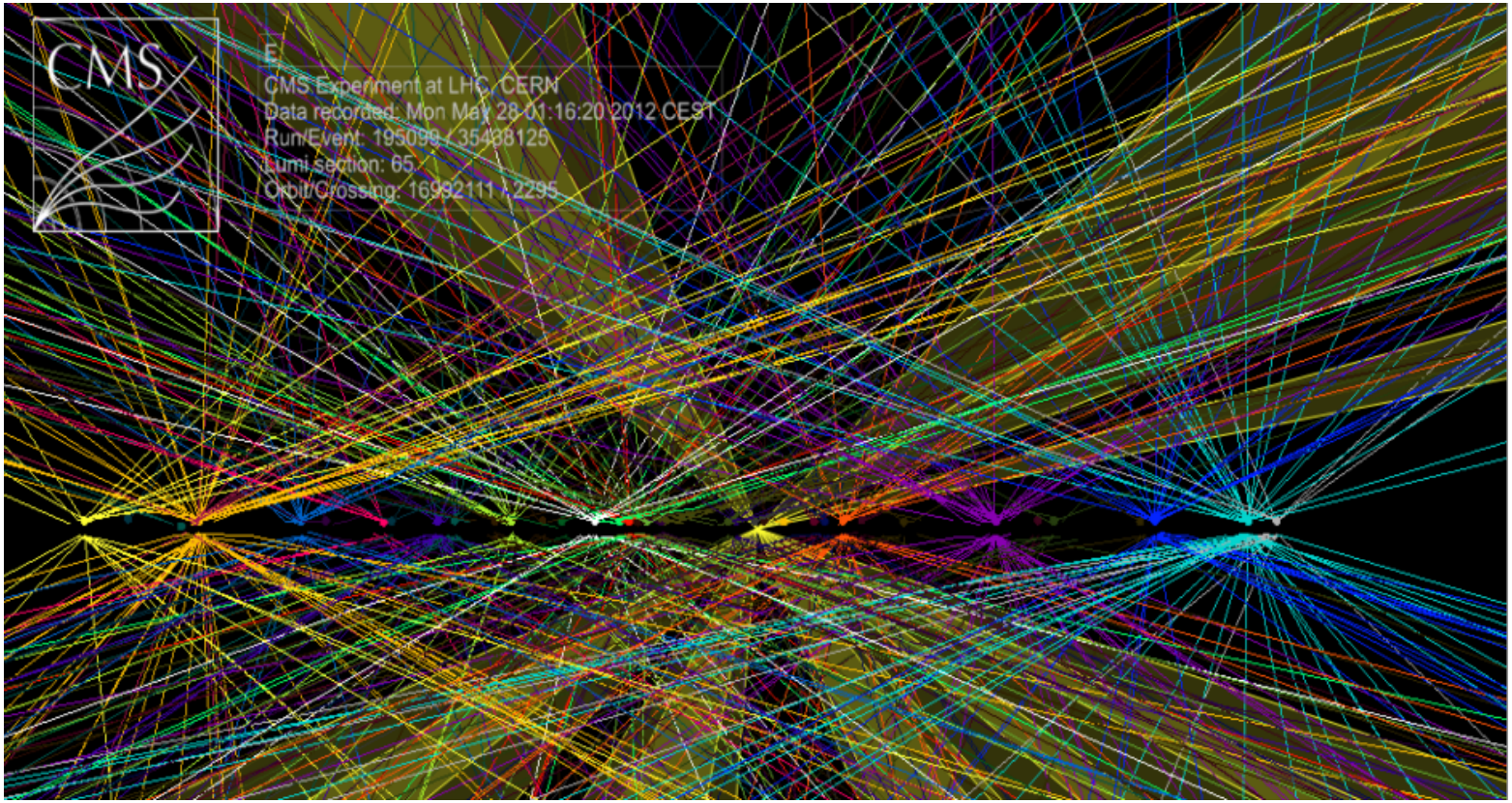


Z->μμ event;
2011 data.

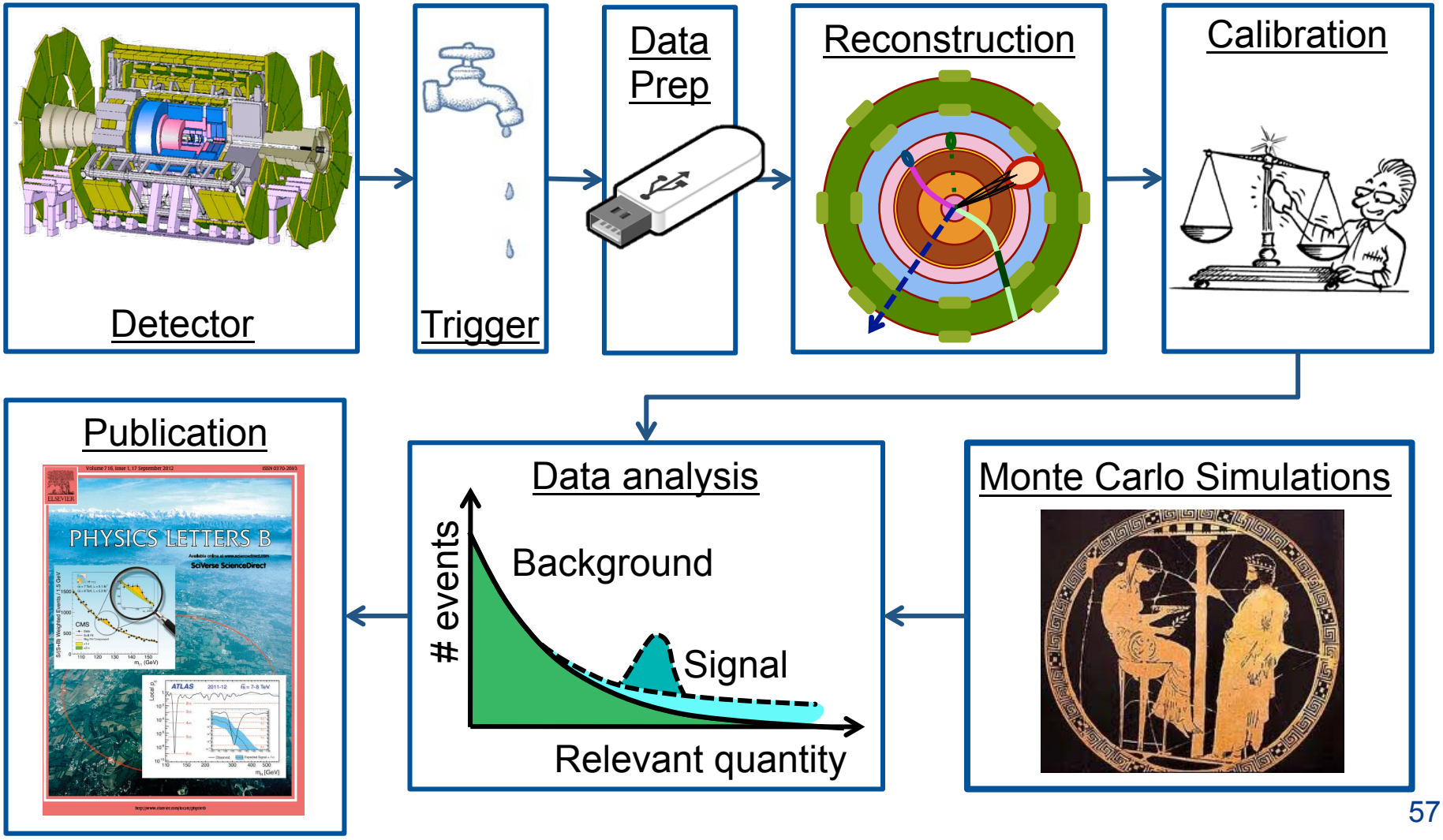
Track pT > 10 GeV

INT / XING

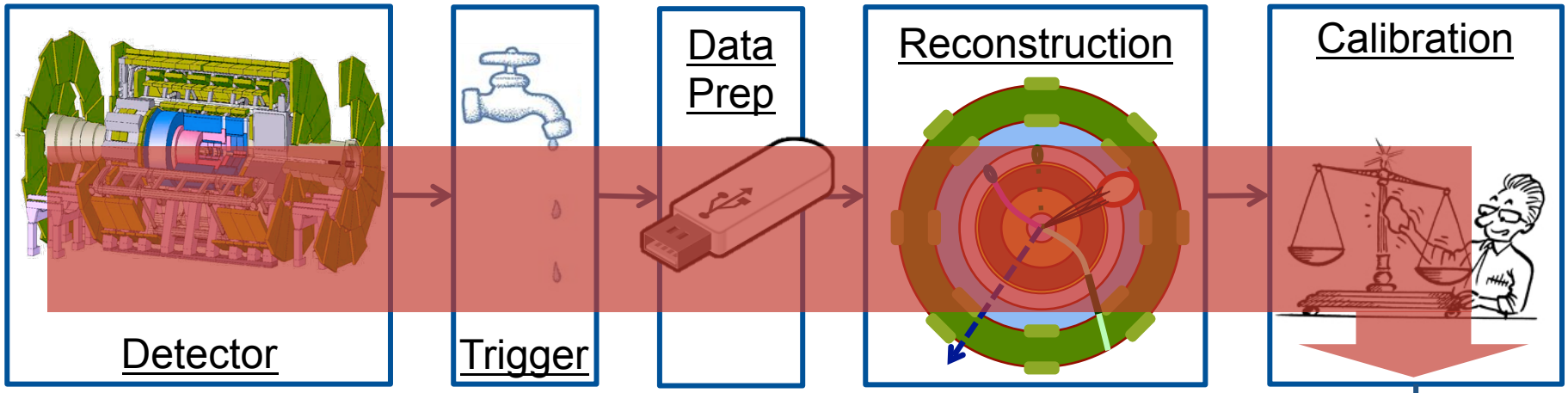

Peak interactions per crossing
50
45
40
35
30
25
20
15
10
5
0
ATLAS
Online Luminosity
√s = 7 TeV
√s = 7 TeV
√s = 8 TeV
Jan
Apr
Jul
Oct
Jan
Apr
Jul
Oct
Jan
Apr
Jul
Oct
Month in 2010
Month in 2011
Month in 2012

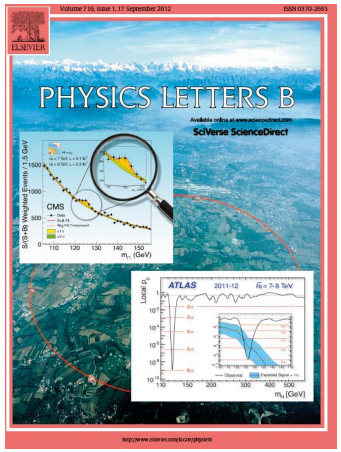

CMS
CMS Experiment at LHC, CERN
Data recorded: Mon May 28 01:16:20 2012 CEST
Run/Event: 195099 / 35488125
Lumi section: 65
Orbit/Crossing: 16932111 / 2295

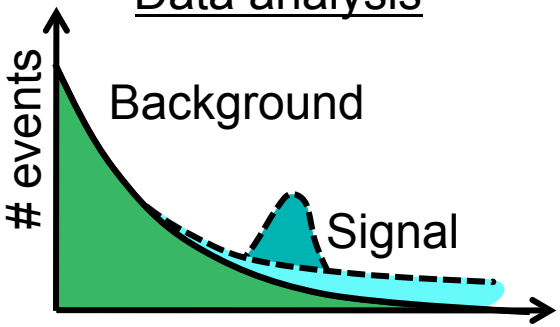
AN EVENT'S LIFETIME



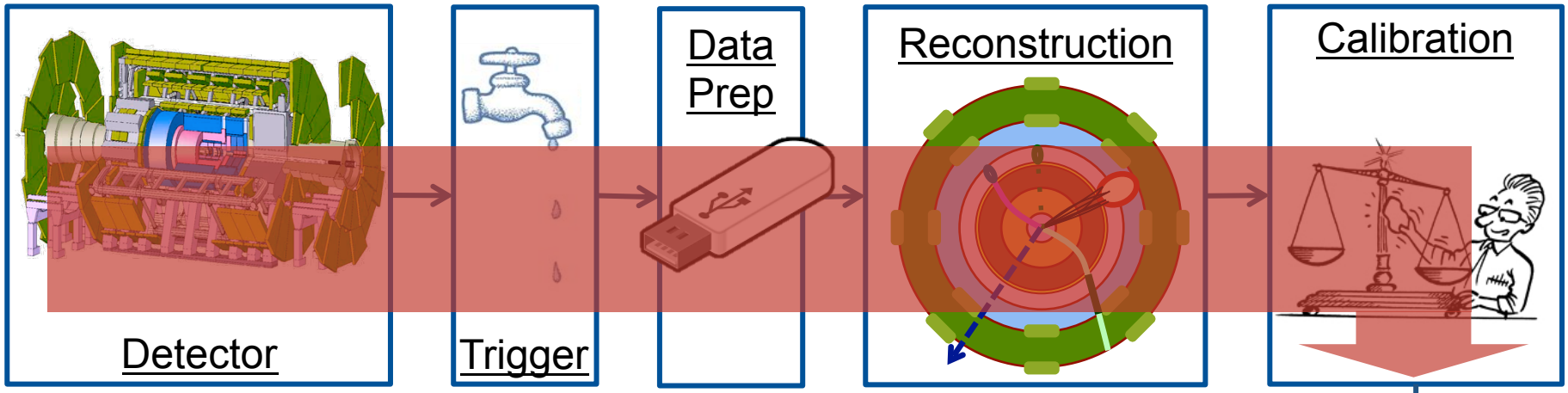
AN EVENT'S LIFETIME

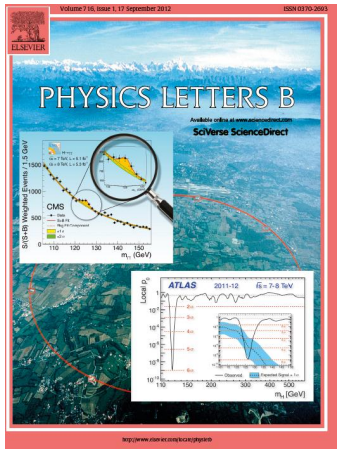


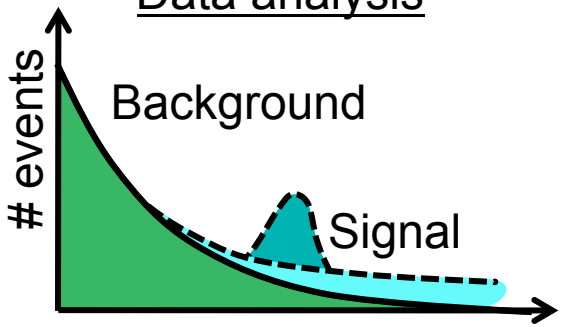




AN EVENT'S LIFETIME



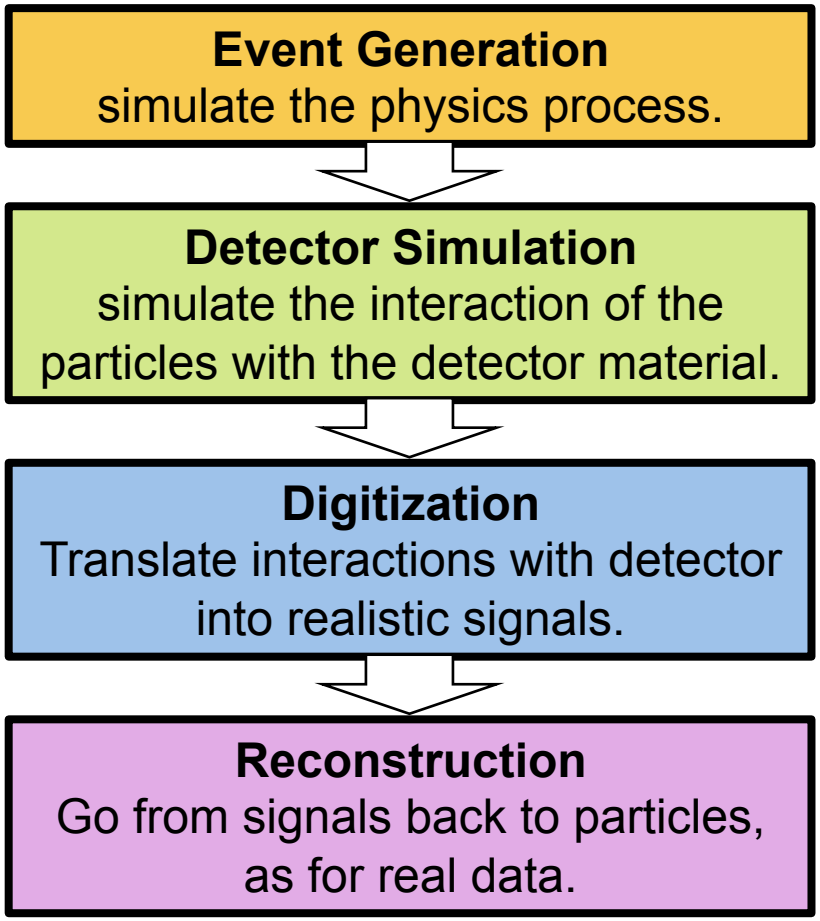




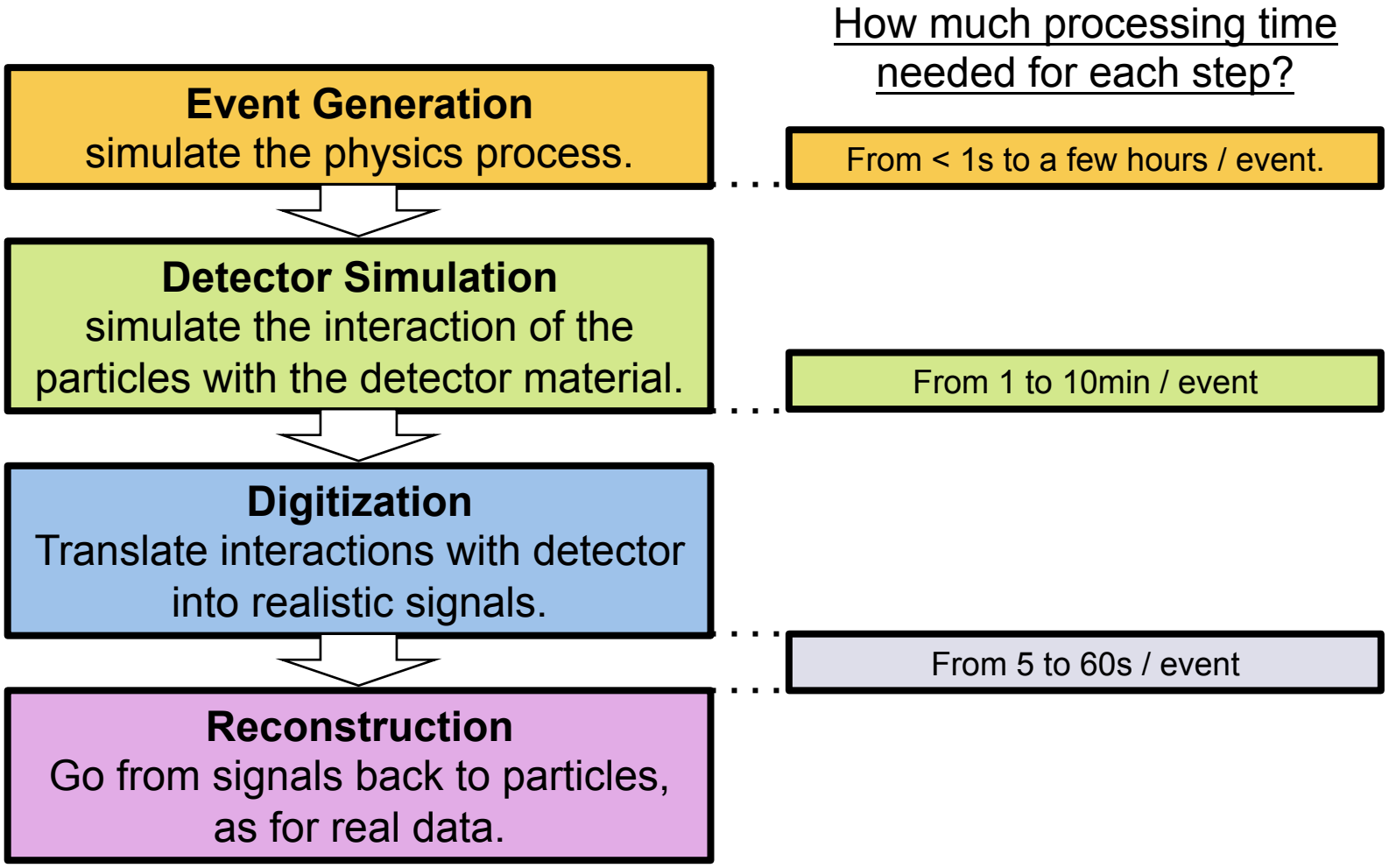
MONTE CARLO SIMULATION – WHY

- **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.**

MONTE CARLO PRODUCTION CHAIN



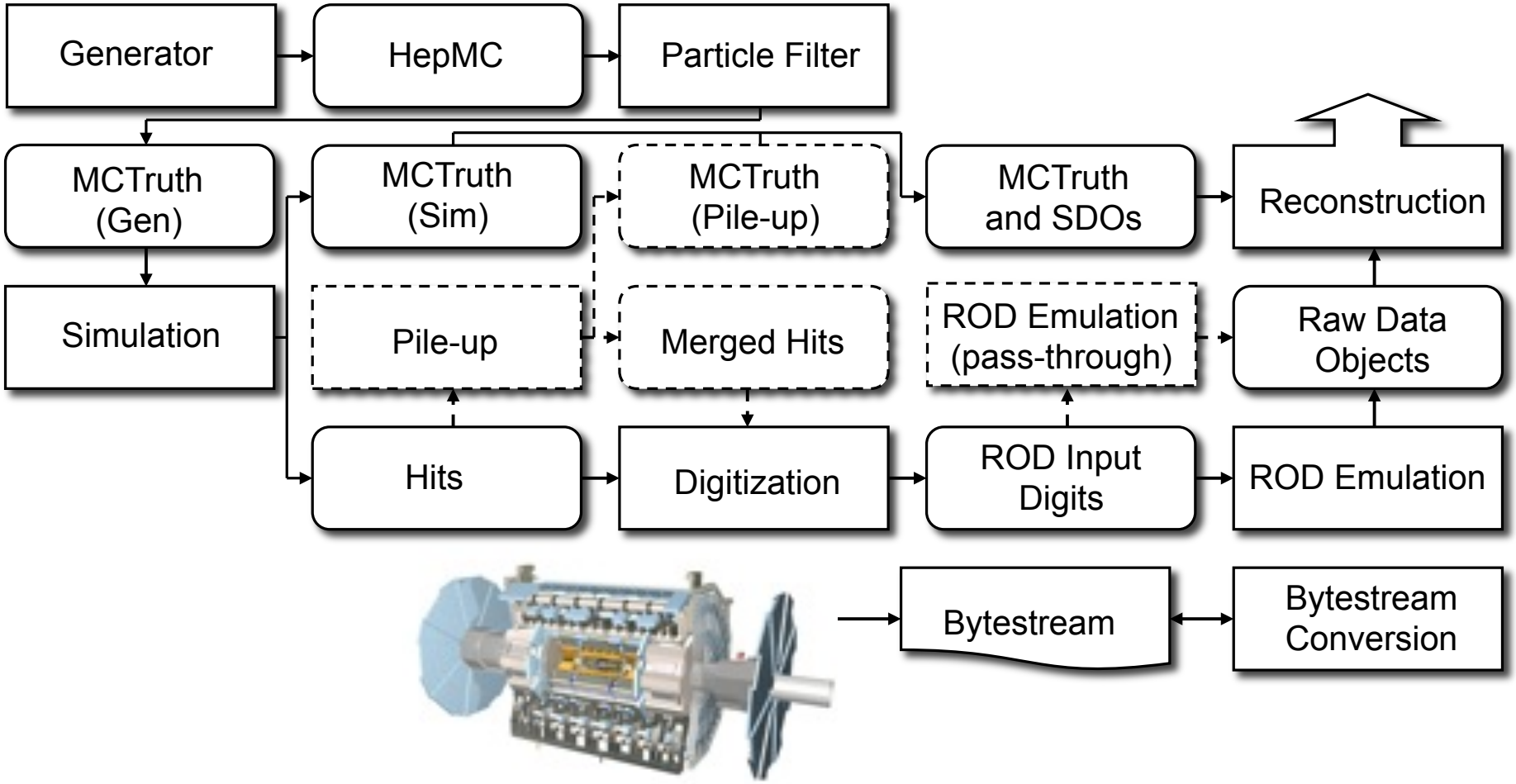
MONTE CARLO PRODUCTION CHAIN



MONTE CARLO PRODUCTION CHAIN

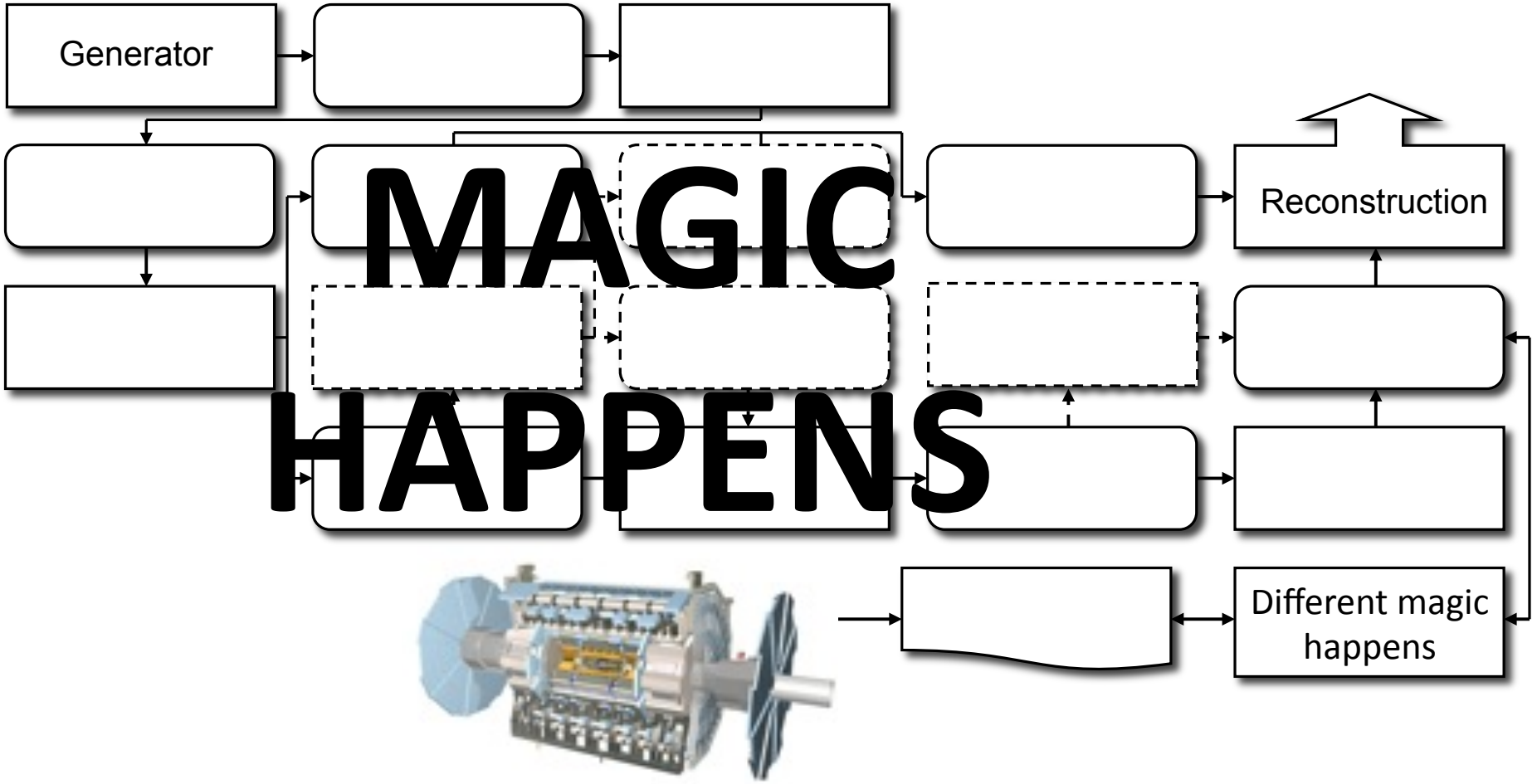
Courtesy: Z. Marshall

Our LHC Simulation: The Dream

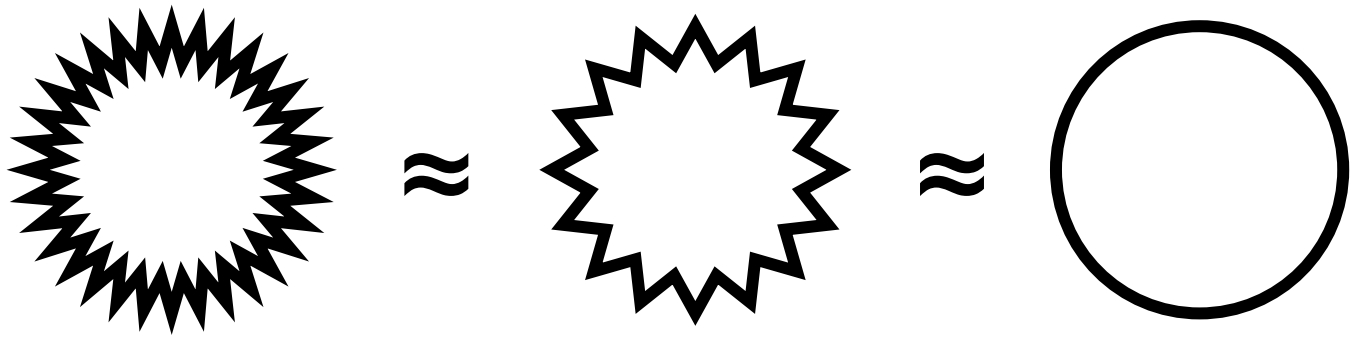


Our LHC Simulation: The Reality?

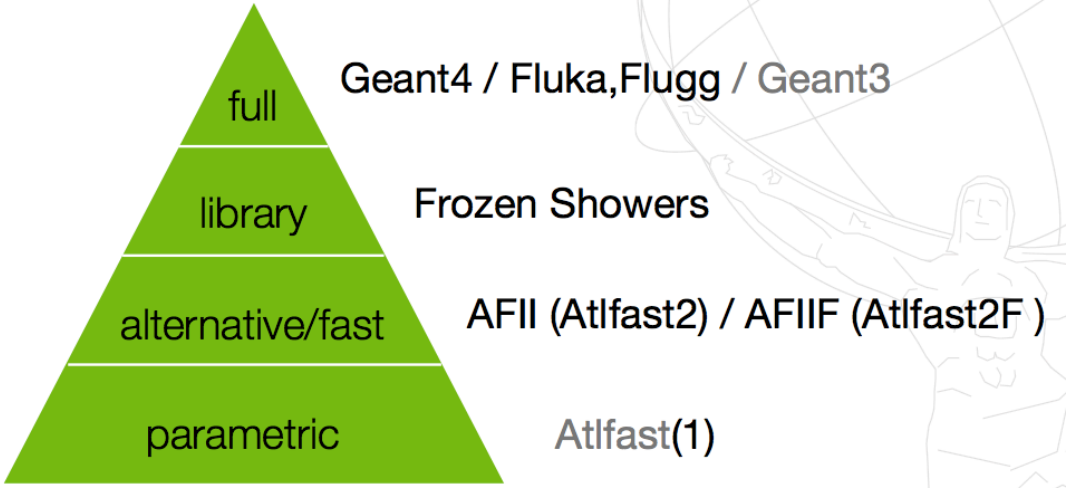
This is most people's view of the chain

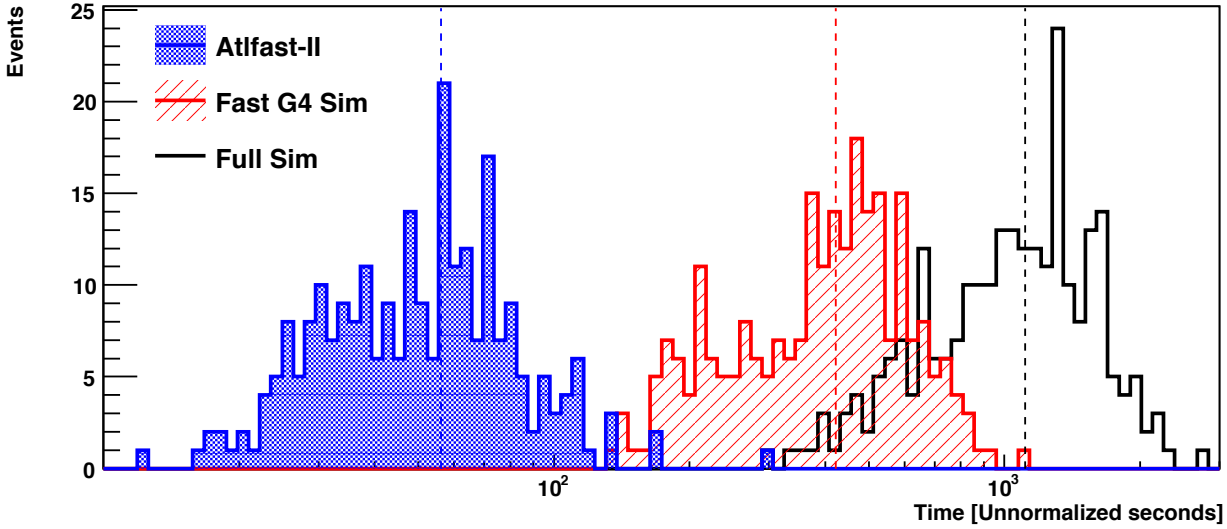


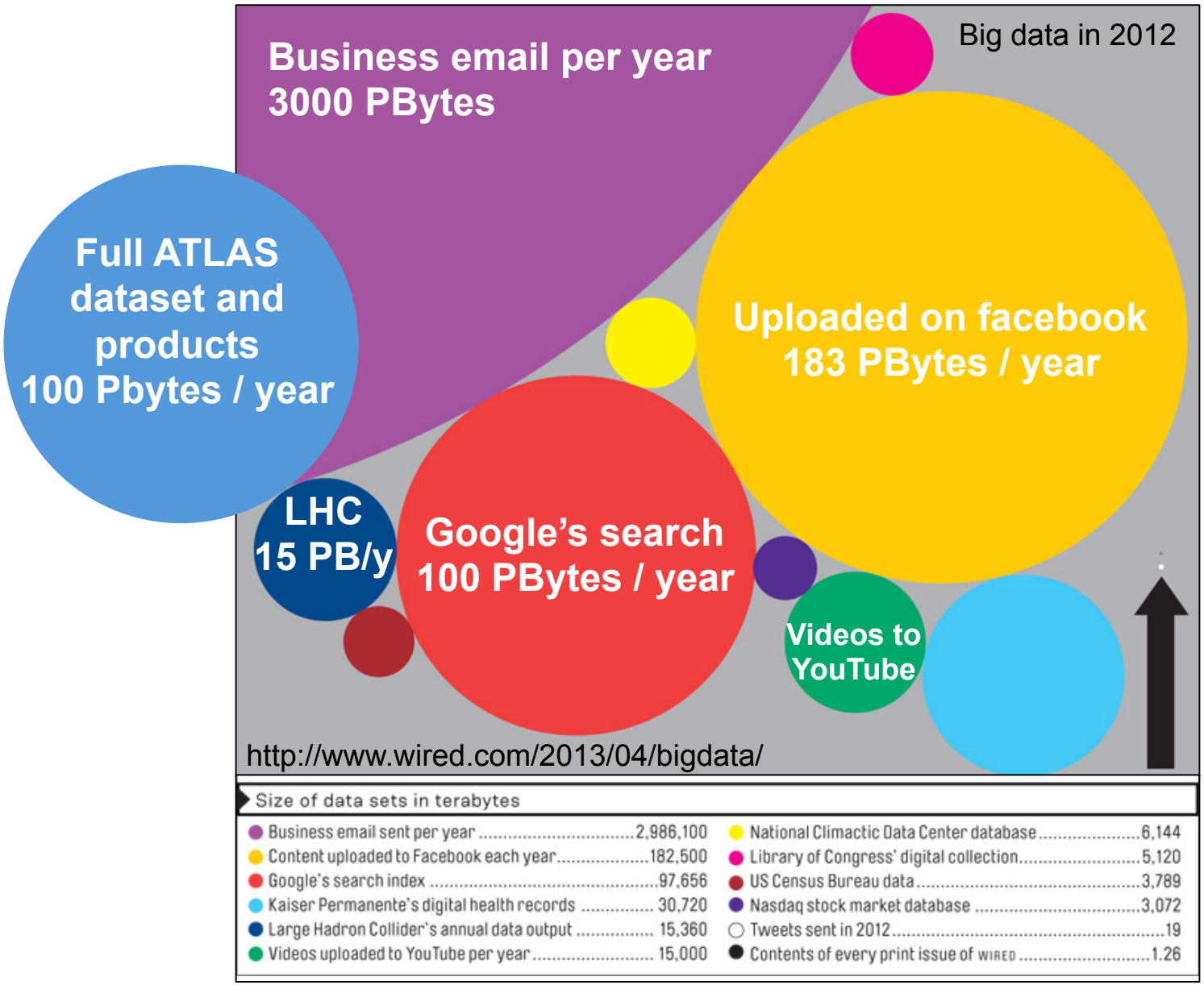
SIMULATION – FULL AND FAST



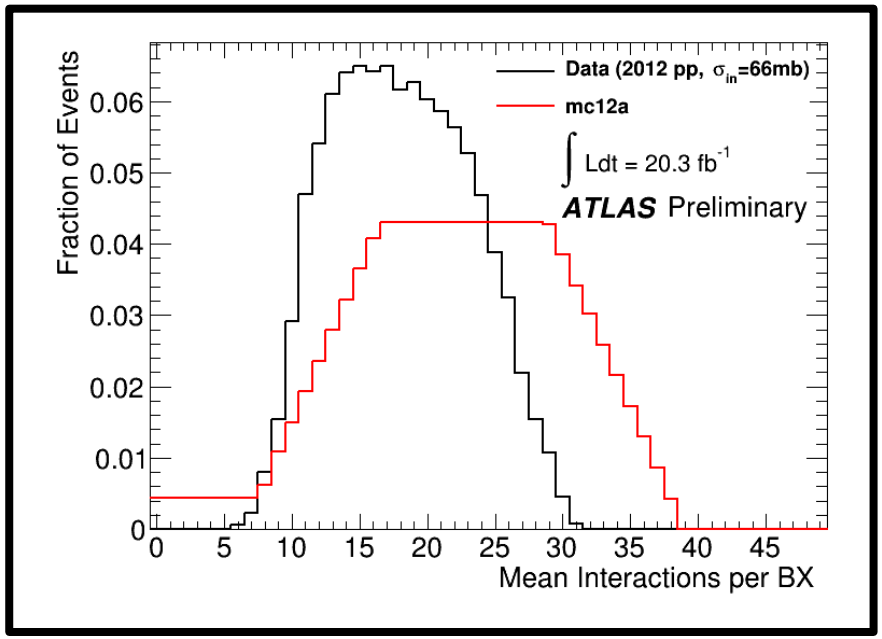
SIMULATION – FULL AND FAST

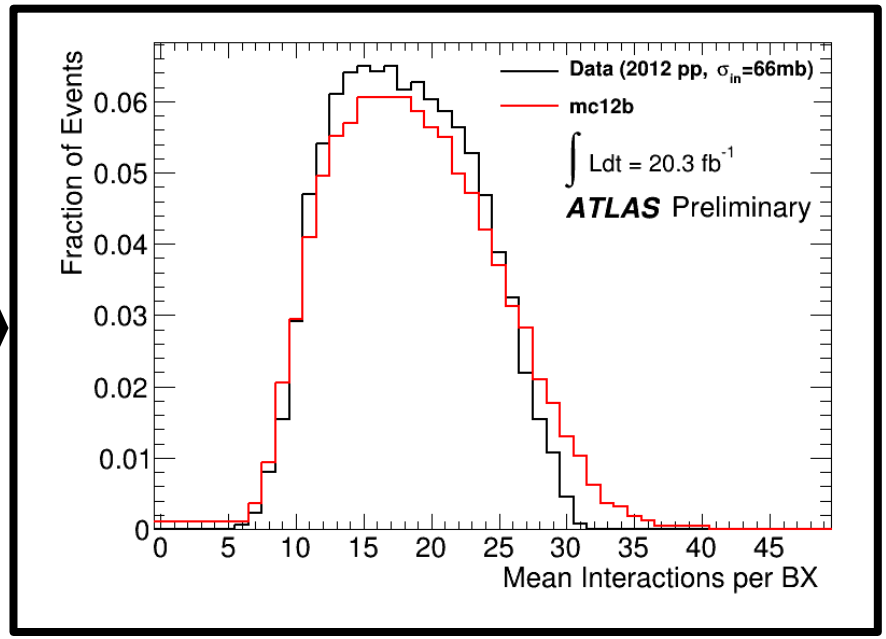




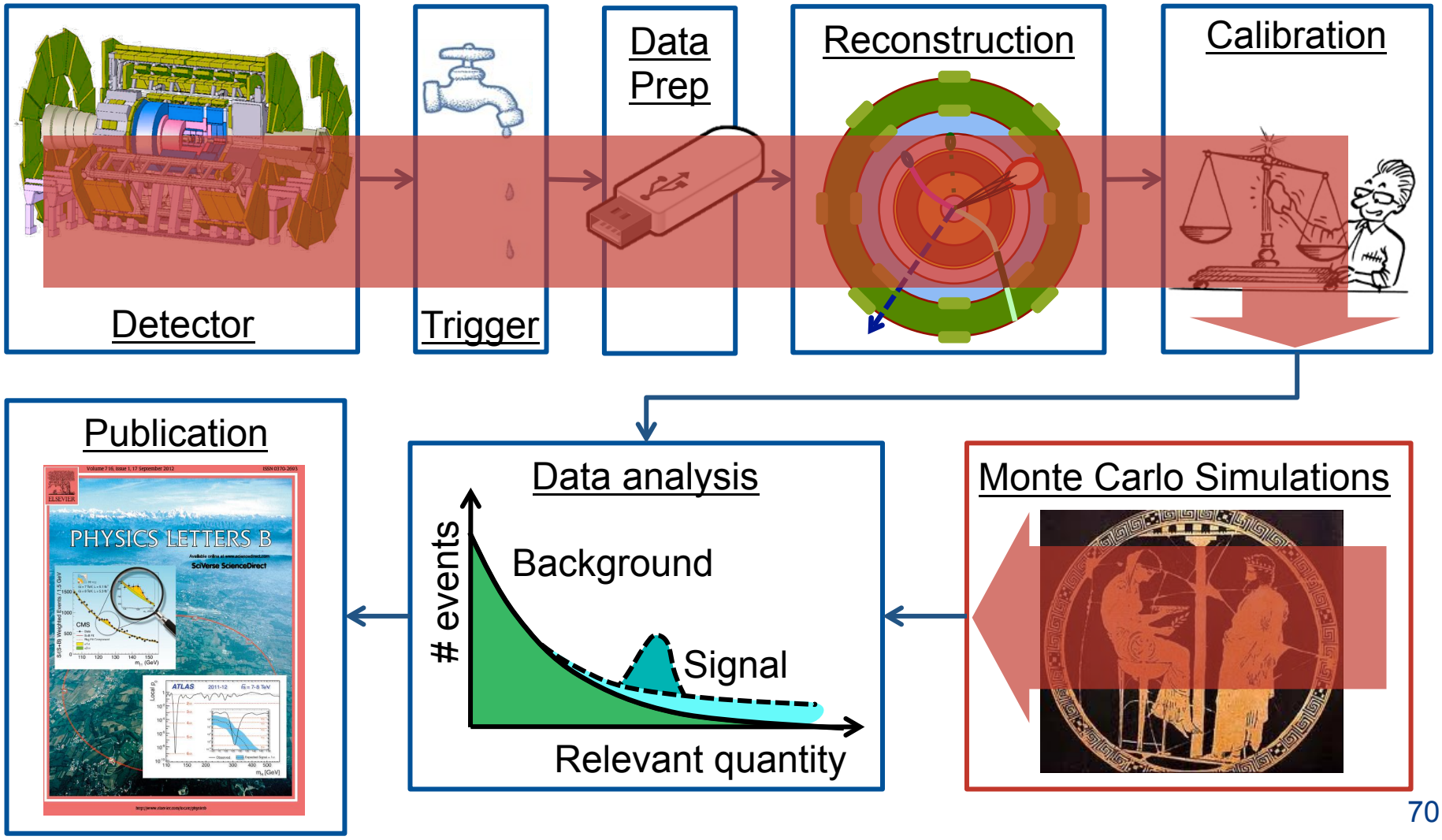

Big data in 2012
Business email per year
3000 PBytes
Full ATLAS
dataset and
products
100 Pbytes / year
Uploaded on facebook
183 PBytes / year
LHC
15 PB/y
Google's search
100 PBytes / year
Videos to
YouTube
http://www.wired.com/2013/04/bigdata/
Size of data sets in terabytes
Business email sent per year 2,986,100
Content uploaded to Facebook each year 182,500
Google's search index 97,656
Kaiser Permanente's digital health records 30,720
Large Hadron Collider's annual data output 15,360
Videos uploaded to YouTube per year 15,000
National Climactic Data Center database 6,144
Library of Congress' digital collection 5,120
US Census Bureau data 3,789
Nasdaq stock market database 3,072
Tweets sent in 2012 19
Contents of every print issue of WIRED 1.26

PILEUP IN SIMULATION





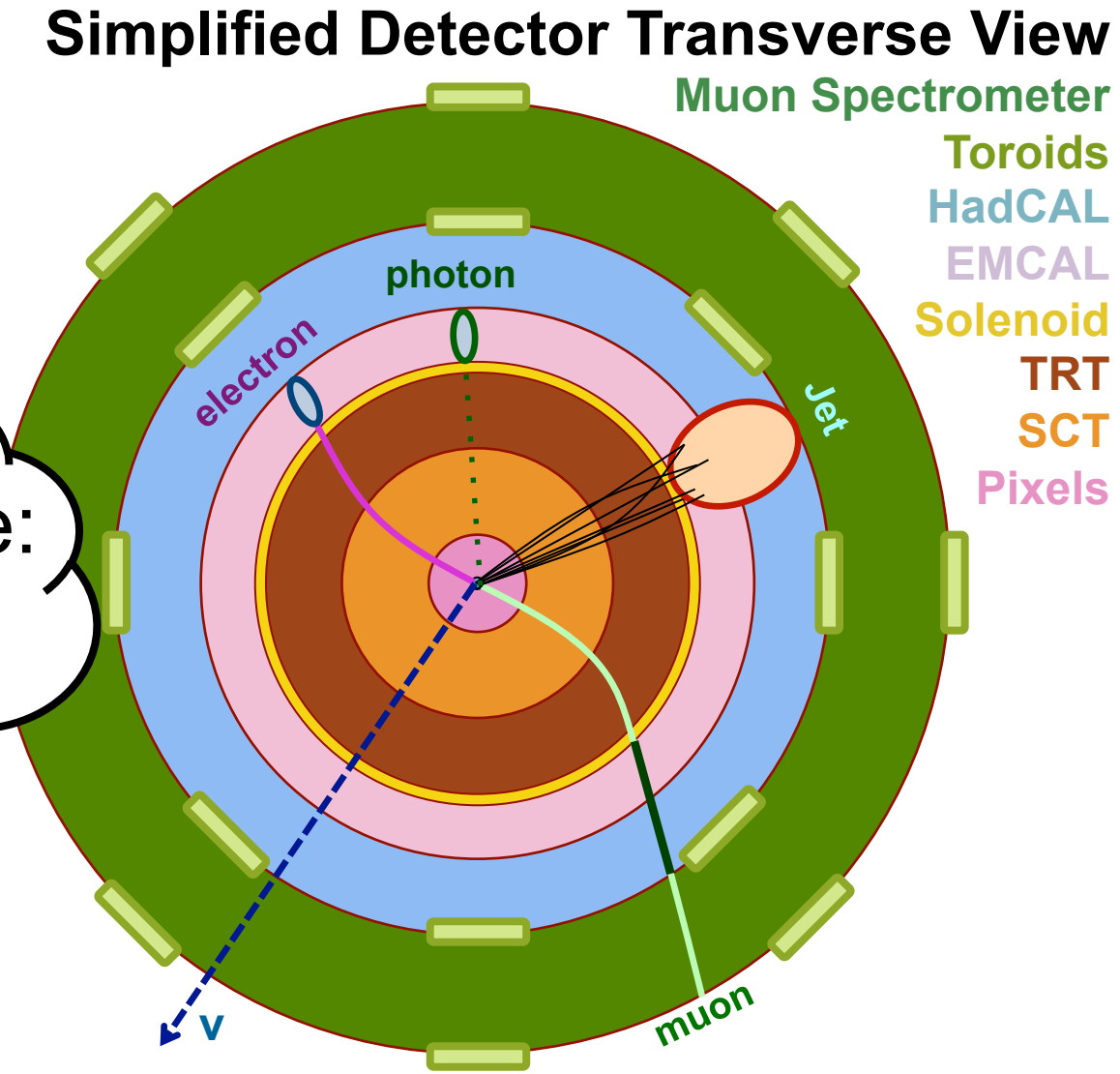
END OF LECTURE 1



WHAT DO WE RECONSTRUCT

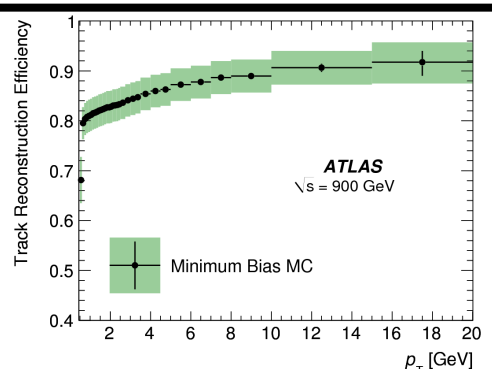
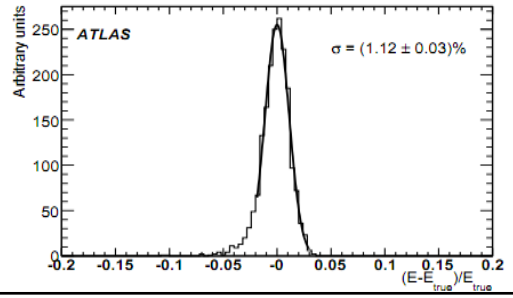
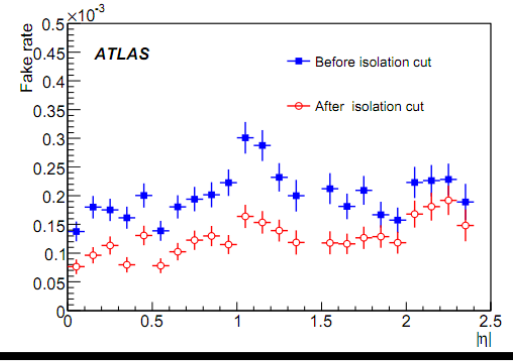
Tracks and Clusters

Combining those: "objects" ("particles")



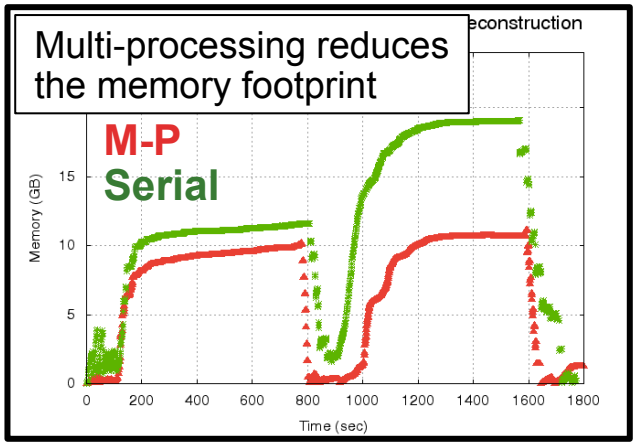
RECONSTRUCTION – FIGURES OF MERIT

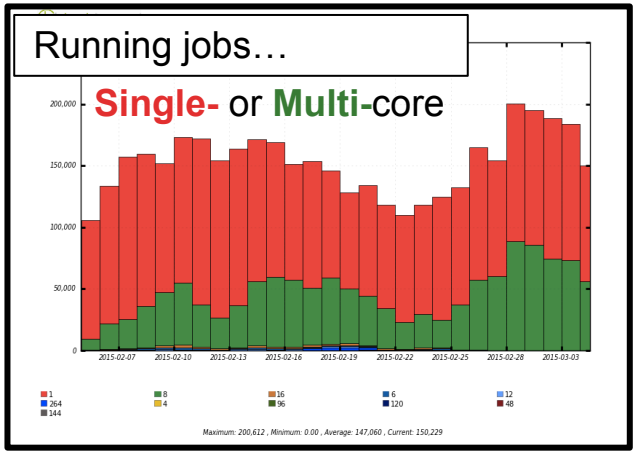
"**true**" quantity: quantity at MC generator level.

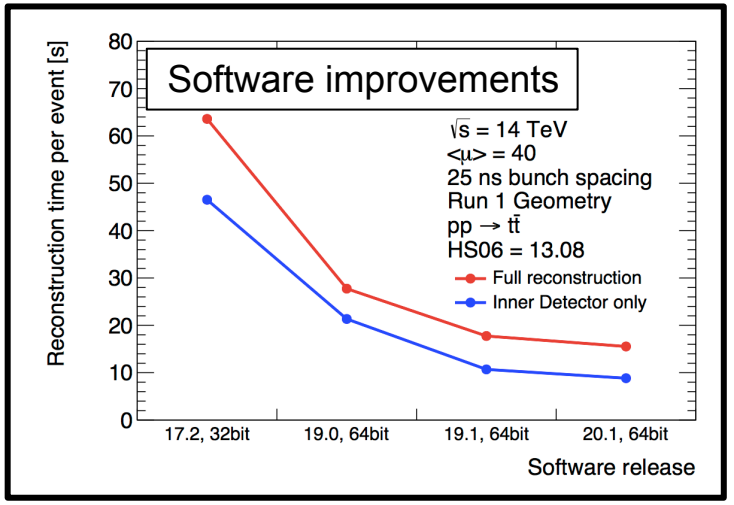
	Definition	Example		Needs be:
Efficiency	how often do we reconstruct the object	tracking efficiency = (number of reconstructed tracks) / (number of true tracks)		High
Resolution	how accurately do we reconstruct the quantity	energy resolution = (measured energy – true energy) / (true energy)		Good
Fake rate	how often we reconstruct a different object as the object we are interested in	a jet faking an electron, fake rate = (Number of jets reconstructed as an electron) / (Number of jets)		Low

RECONSTRUCTION – GOALS

- **High efficiency.**
- **Good resolution.**
- **Low fake rate.**
- **Robust against detector problems and data-taking conditions.**
  - Noise.
  - Dead regions of the detector.
  - Increased pile-up.
- **Computing-friendly.** →
  - CPU time per event.
  - Memory use.

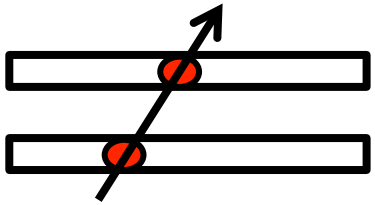




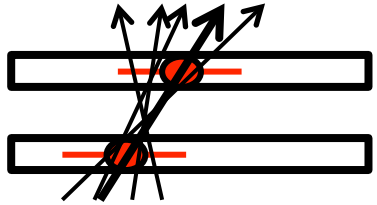


TRACKING IN A NUTSHELL – TRACK FITTING

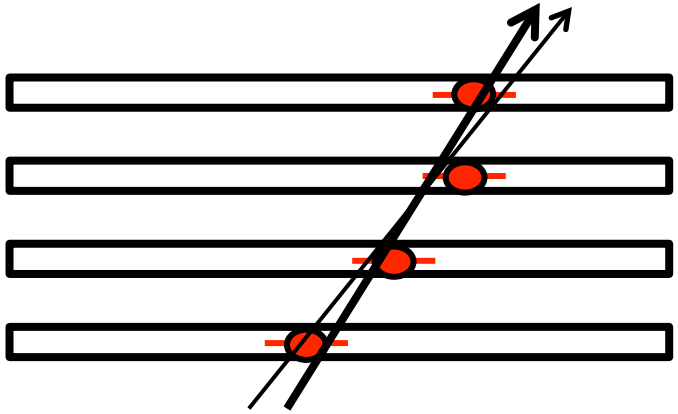
- **Perfect measurement – ideal**



- **Imperfect measurement – reality**



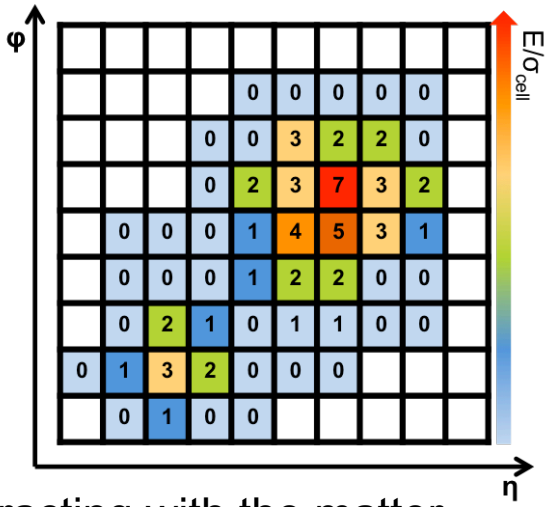
- **Small errors and more points help to constrain the possibilities**



- **Quantitatively:**
  - Parameterize the track;
  - Find parameters by Least-Squares-Minimization;
  - Obtain also uncertainties on the track parameters.

CLUSTERING IN A NUTSHELL

- **Reconstruct energy deposited in the calorimeter by charged or neutral particles; electrons, photons and jets.**
- **For a cluster we measure:**
  - The energy;
  - The position of the deposit;
  - The direction of the incident particles;

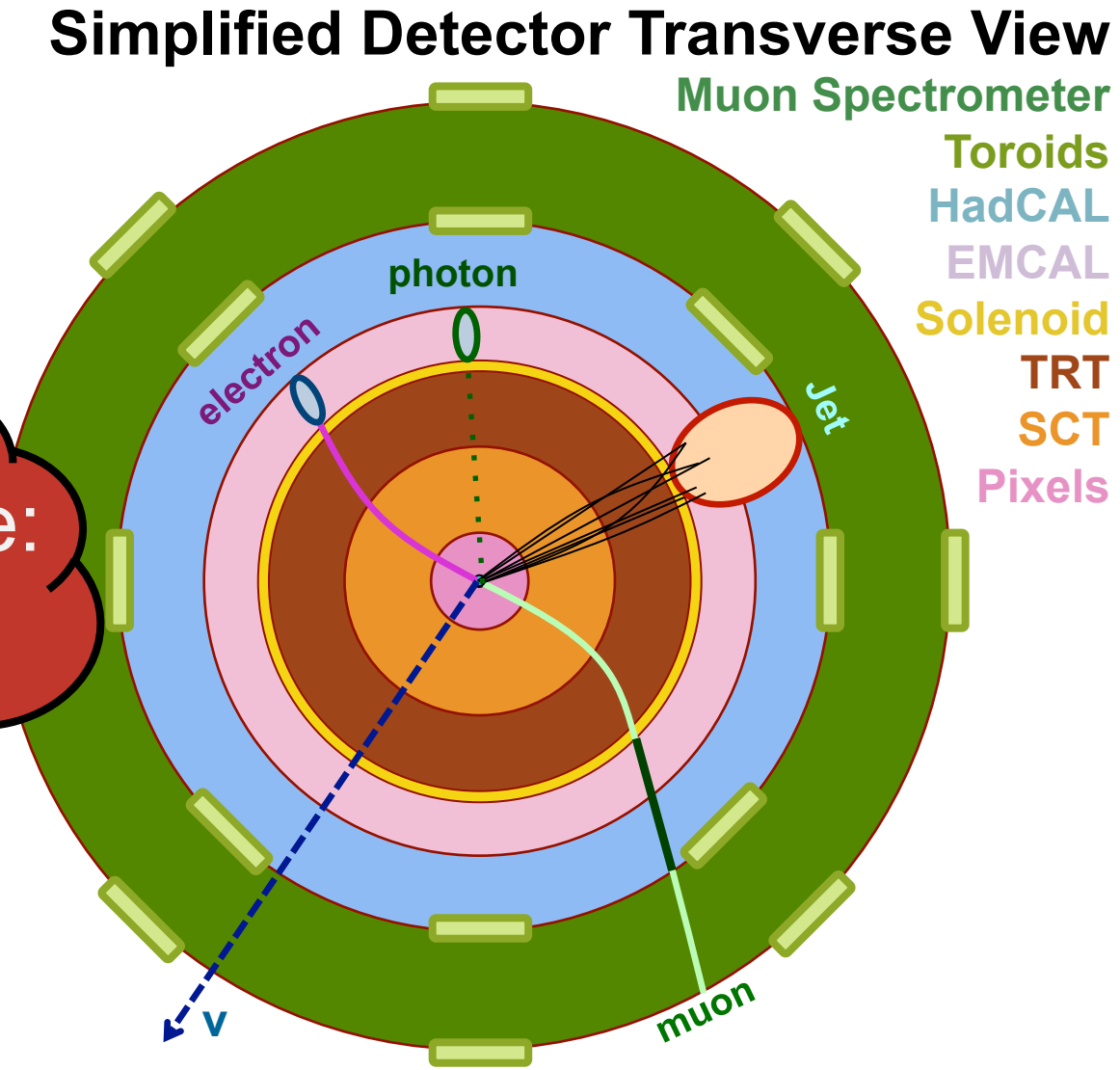


- **Calorimeters are segmented in cells.**
  - Typically a shower created by a particle interacting with the matter extends over several cells.
- **Various clustering algorithms, e.g.:**
  - **Sliding window**. Sum cells within a fixed-size rectangular window.
  - **Topo-clustering**. Start with a seed cell and iteratively add to the cluster the neighbor of a cell already in the cluster.

WHAT DO WE RECONSTRUCT

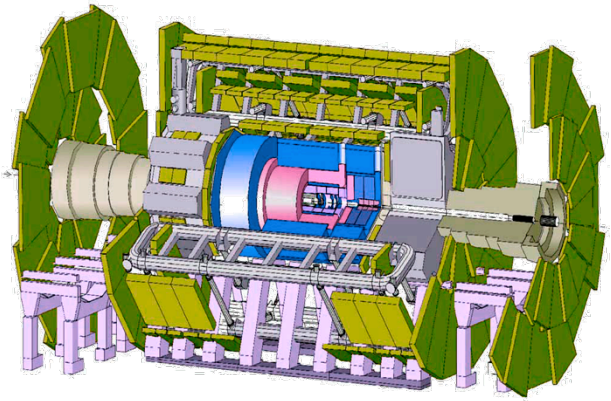
Tracks and Clusters

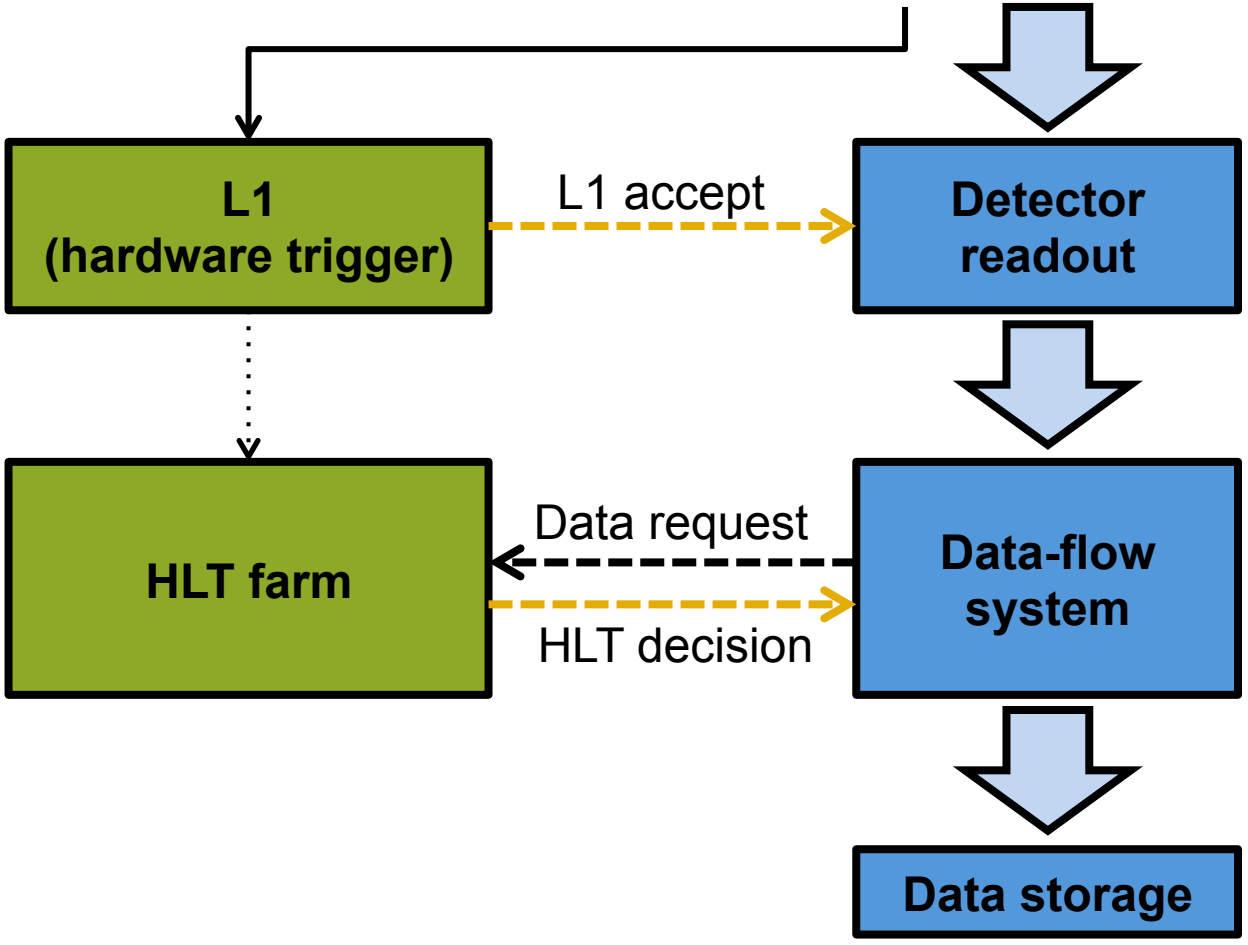
Combining those:
"objects"
("particles")



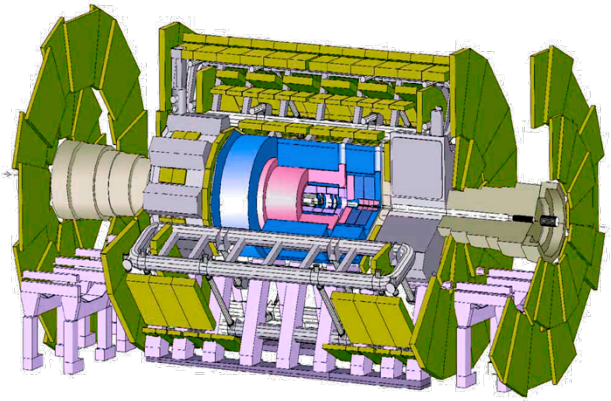
BACKUP

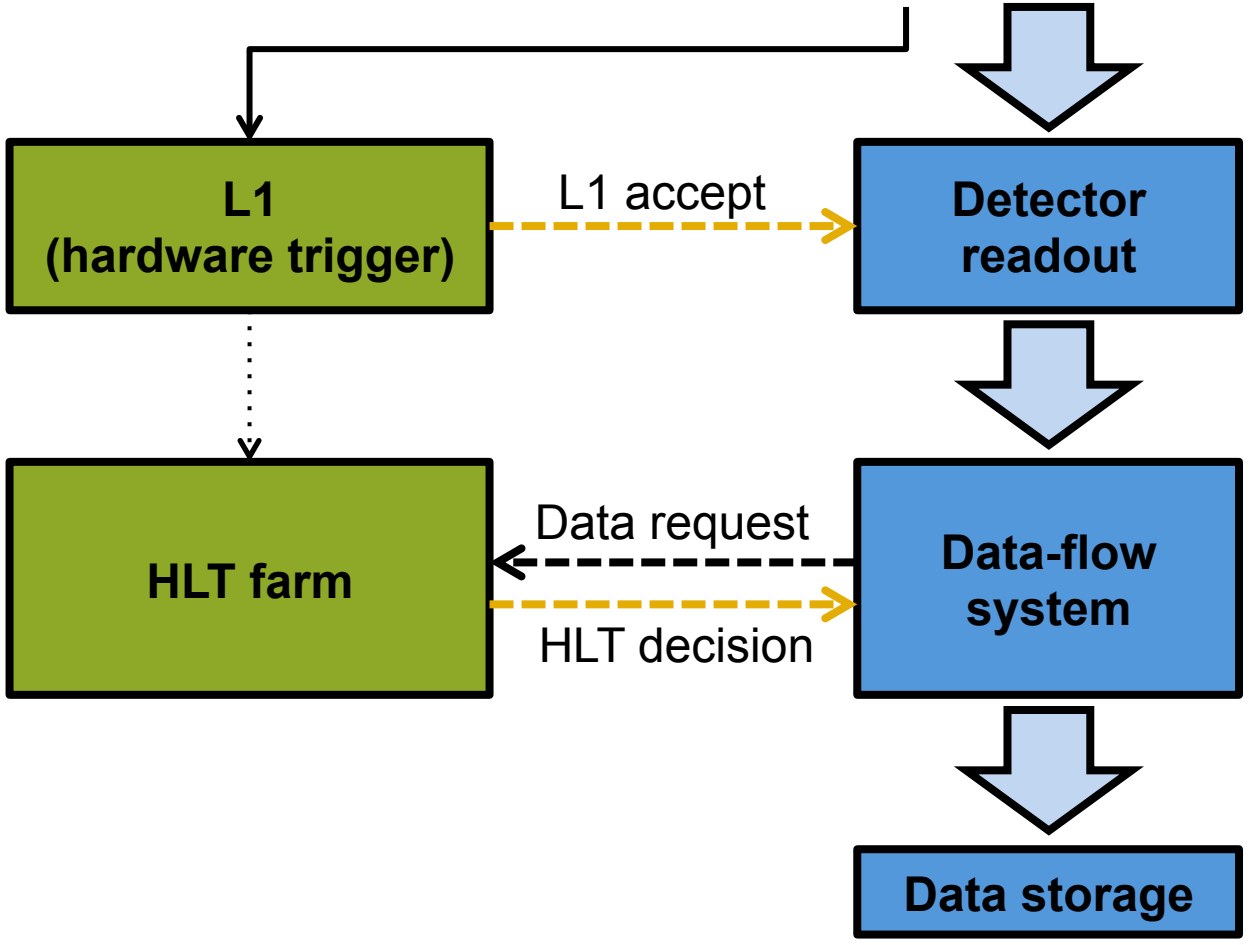
THE DATA ACQUISITION

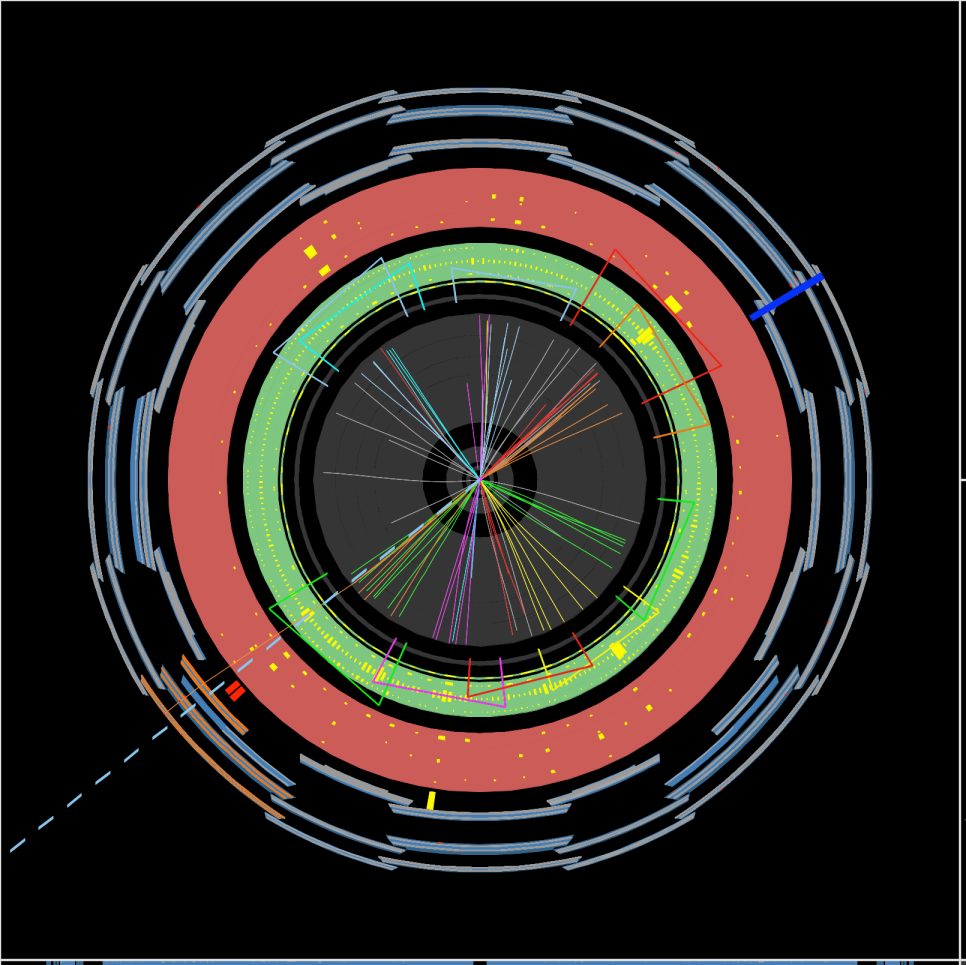


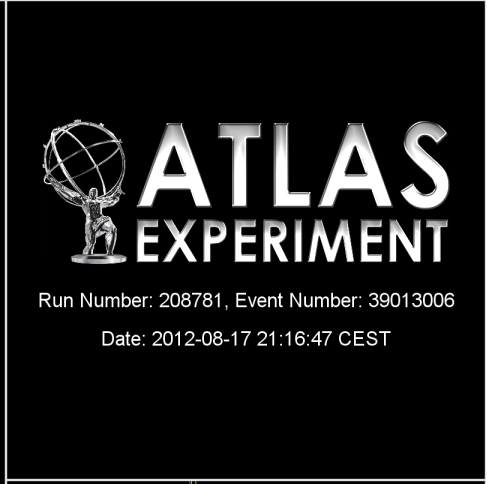


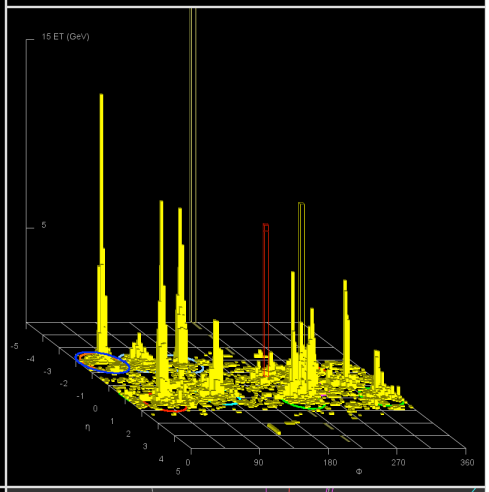
THE DATA ACQUISITION

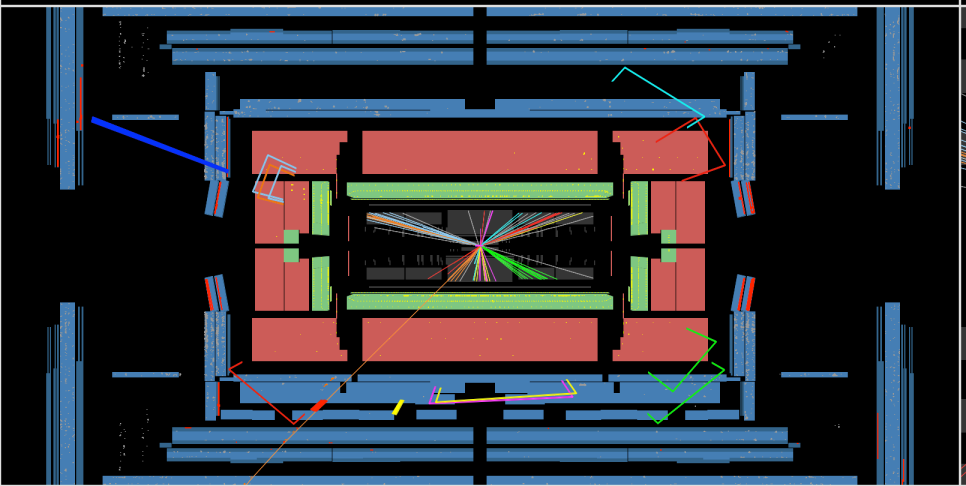


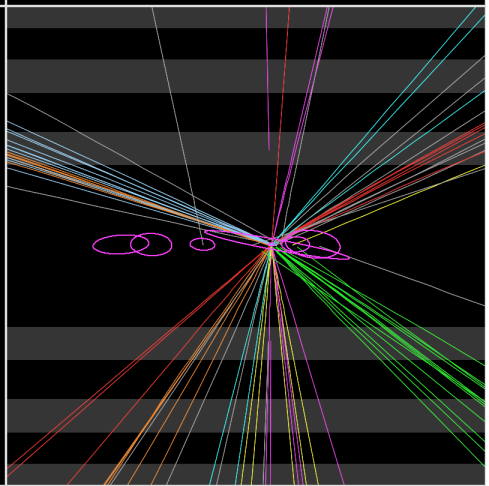






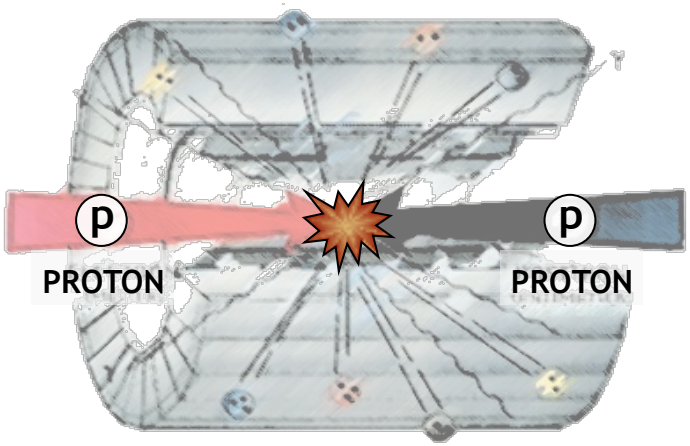


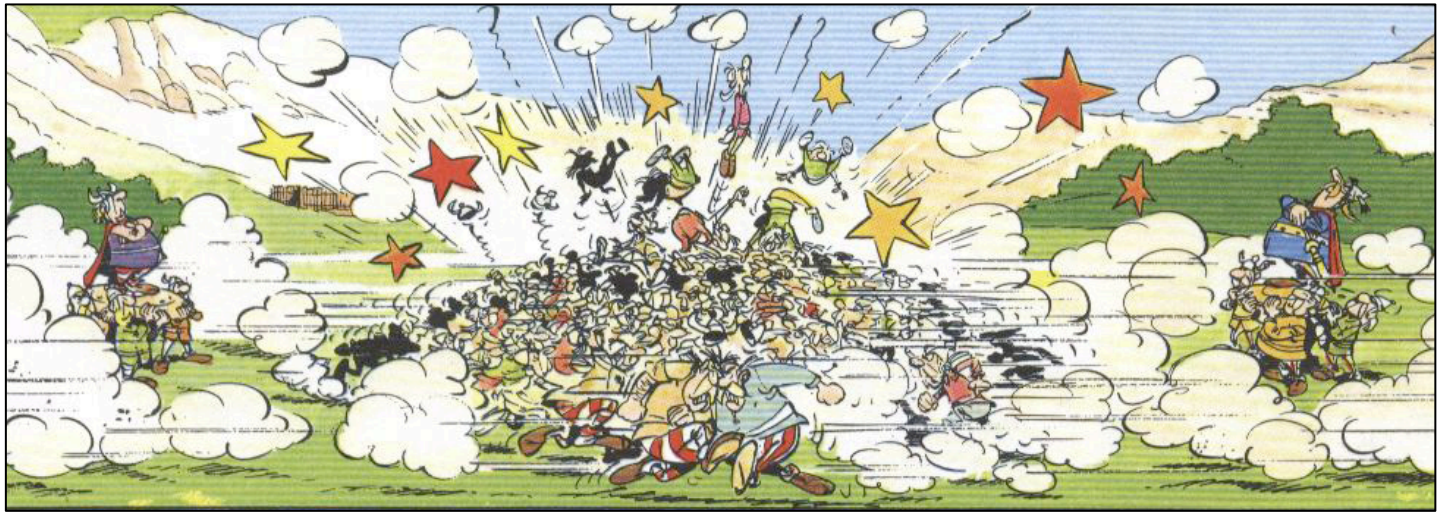




10 jets
with pT > 50GeV
ME_T = 120 GeV

IN A P-P COLLISION





TRIGGER MENUS FOR SUSY

Selection	EF trigger election	EF Avg. Rate (Hz) L_{avrg}=5e33/cm^{2}s
Single jet & E_T^{miss}	Jet E_T >145 GeV & EF-only E_T^{miss} >70 GeV	8
Single jet & E_T^{miss} & $\Delta\phi$(jet,E_T^{miss})	Jet E_T >80 GeV & E_T^{miss} >70 GeV & $\Delta\phi$ >1.0 rad	8
H_T	>700 GeV	8
Single electron & E_T^{miss}	Electron p_T >25 GeV & EF-only E_T^{miss} >35 GeV	26
Single muon & single jet & E_T^{miss}	Muon p_T >24 GeV & jet E_T >65 GeV & EF-only E_T^{miss} >40 GeV	15
Single photon & E_T^{miss}	Photon p_T >40 GeV & EF-only E_T^{miss} >60 GeV	5
3 electrons	p_T > 18, 2×7 GeV	<1
3 muons	p_T > 18, 2×4 GeV	<1
3 electrons & muons	p_T > 2×7 (e), 6 (μ) GeV	<1
	p_T > 7 (e), 2×6 (μ) GeV	<1

‘DELAYED’ TRIGGERS

Trigger	EF trigger Selection	
	Prompt Stream	Delayed Stream
Multi-jets	4×80 GeV	4×65 GeV
	5×55 GeV	5×45 GeV
	6×45 GeV	
H_T	700 GeV	500 GeV
Single jet (R = 1.0)	460 GeV	360 GeV
E_T^{miss}	80 GeV	60 GeV

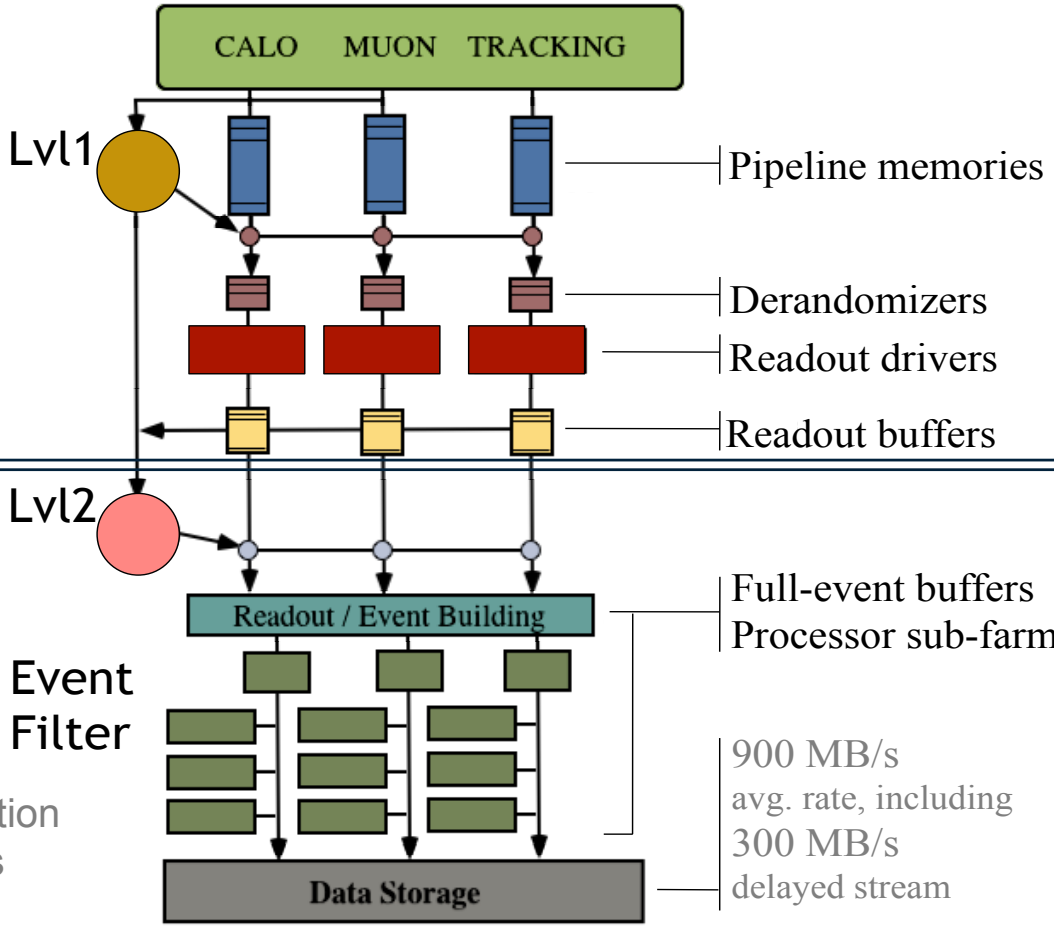
THE ATLAS TRIGGER SYSTEM

Rate (2012 conditions)

Hardware

20 MHz
Bunch crossing rate
6.4×10^8
Interactions/s

75 kHz
Peak rate



Software

6 kHz
Peak rate

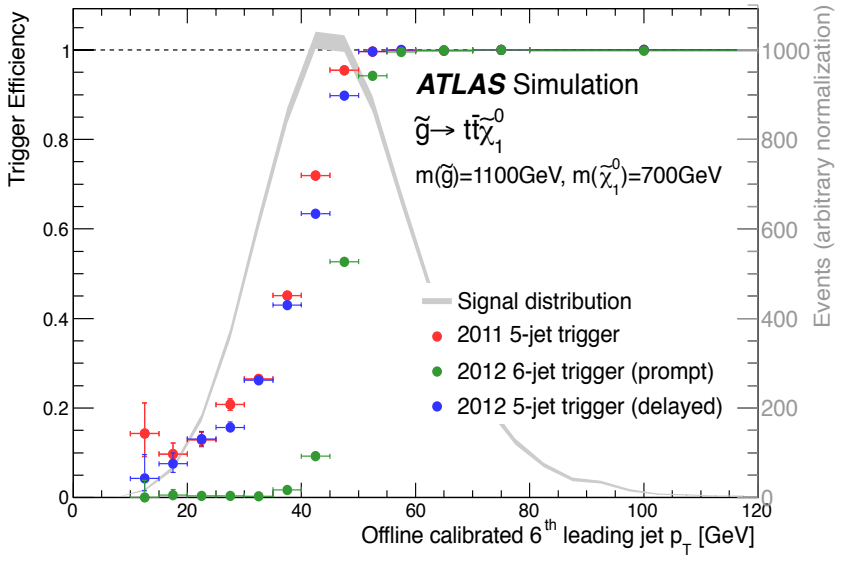
600 Hz
Avg. rate, including
200 Hz delayed stream
(stored for later reconstruction
when computing resources
available)

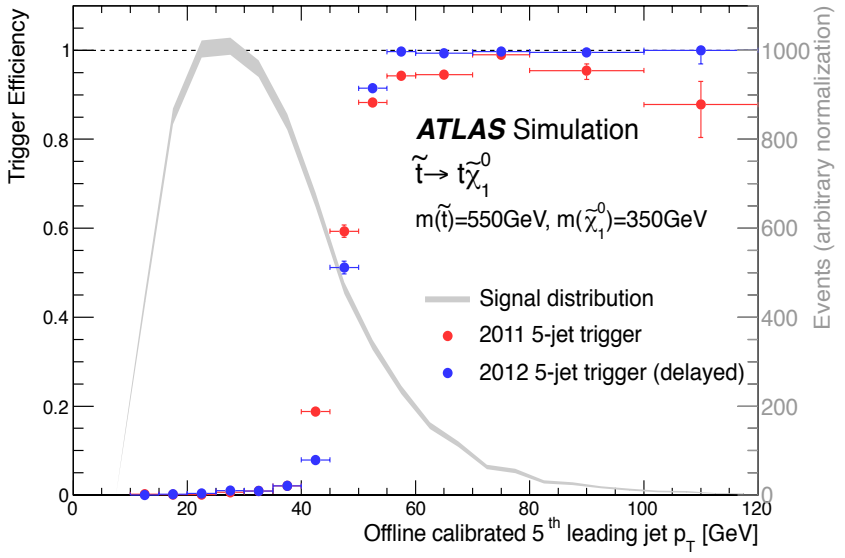
TRIGGER

Signal triggers		
Jet Multiplicity	pT cut	|η|
6	45	3.2
5	55	

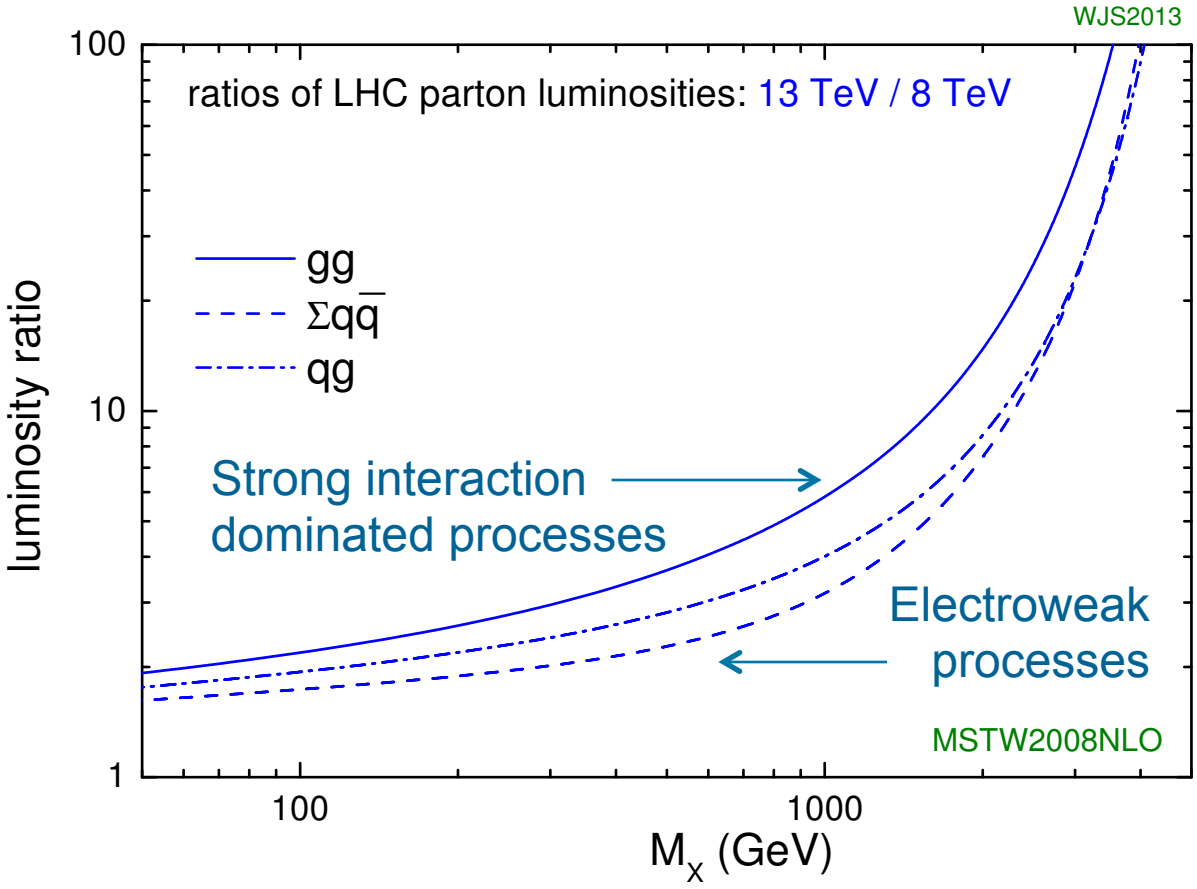
Background/support triggers	
Type	Purpose
Multijet (prescaled)	Efficiencies & Control regions
Single lepton	Control regions

Multijet trigger improvements in 2012

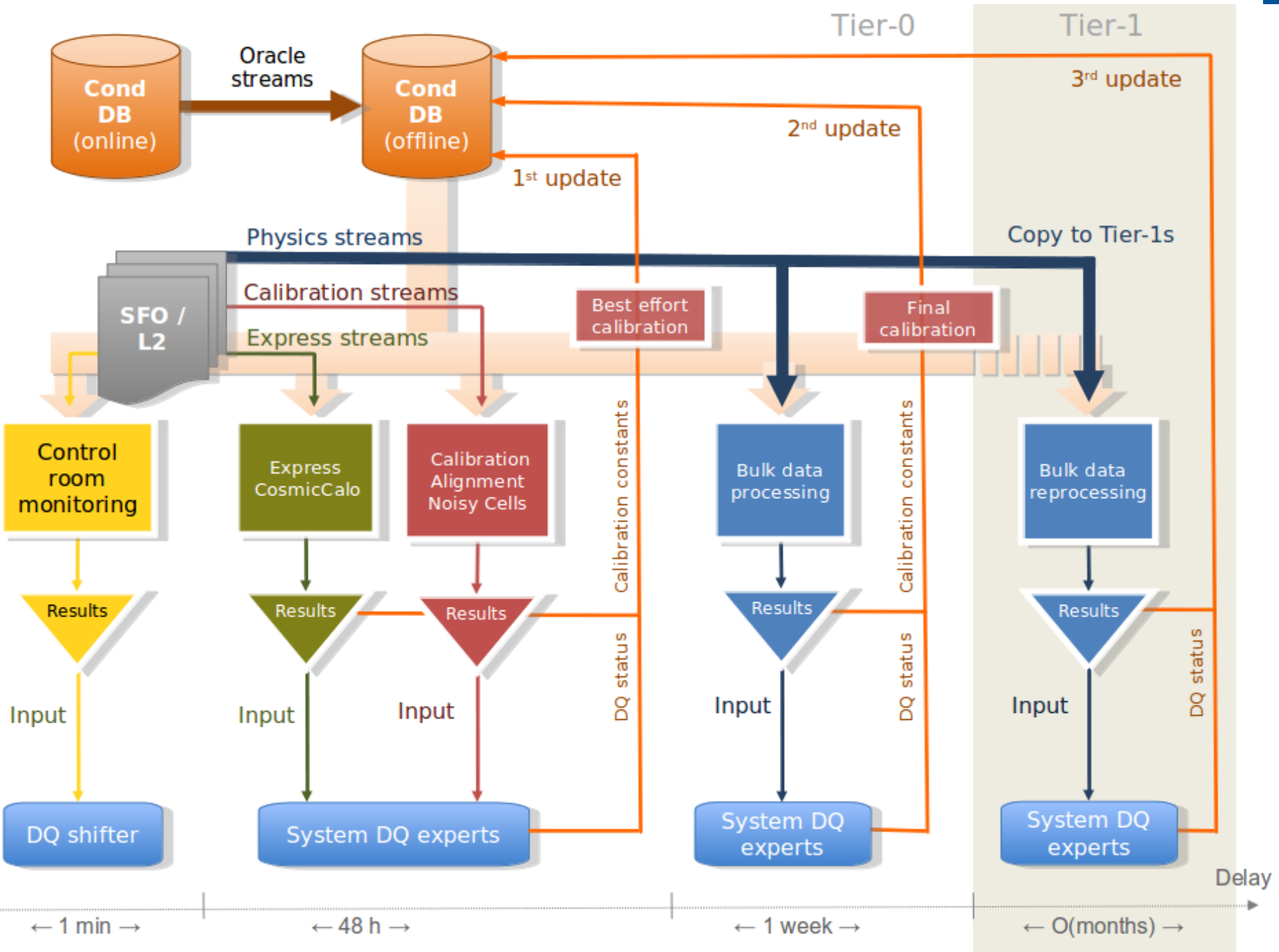




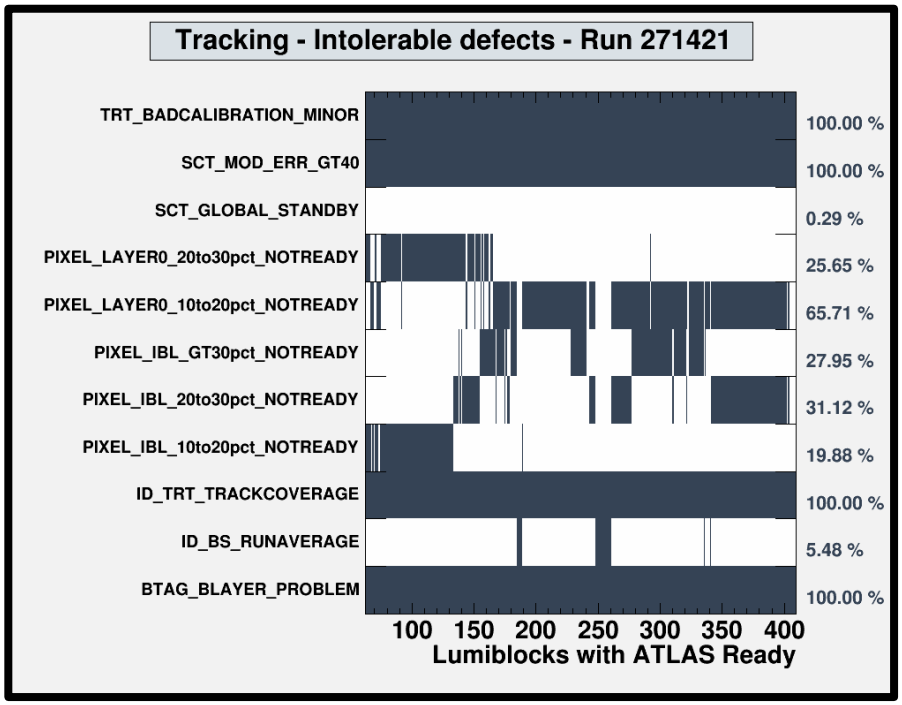
THE BENEFITS

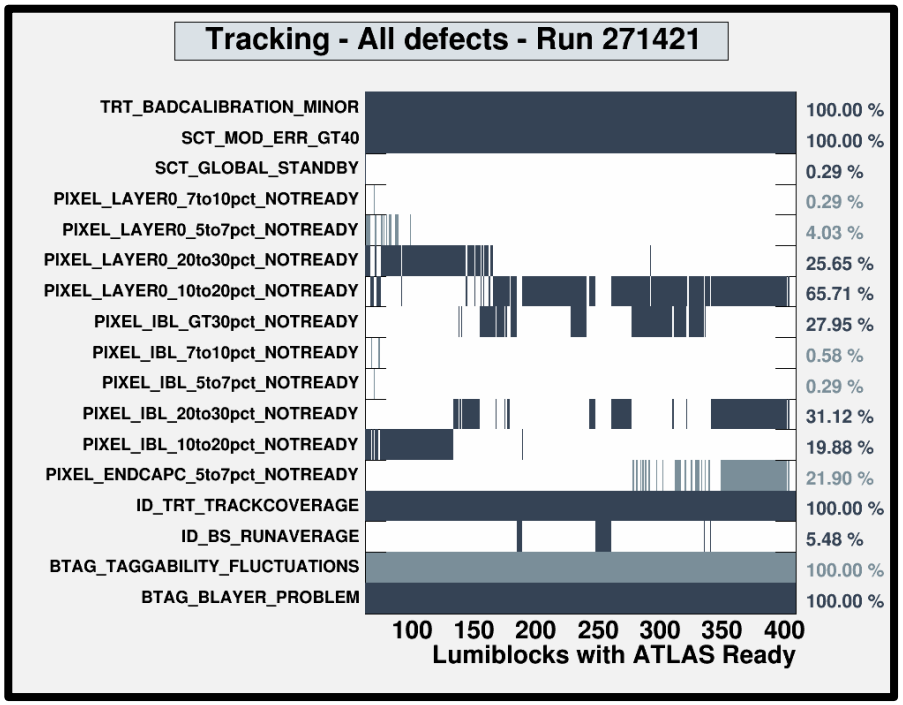


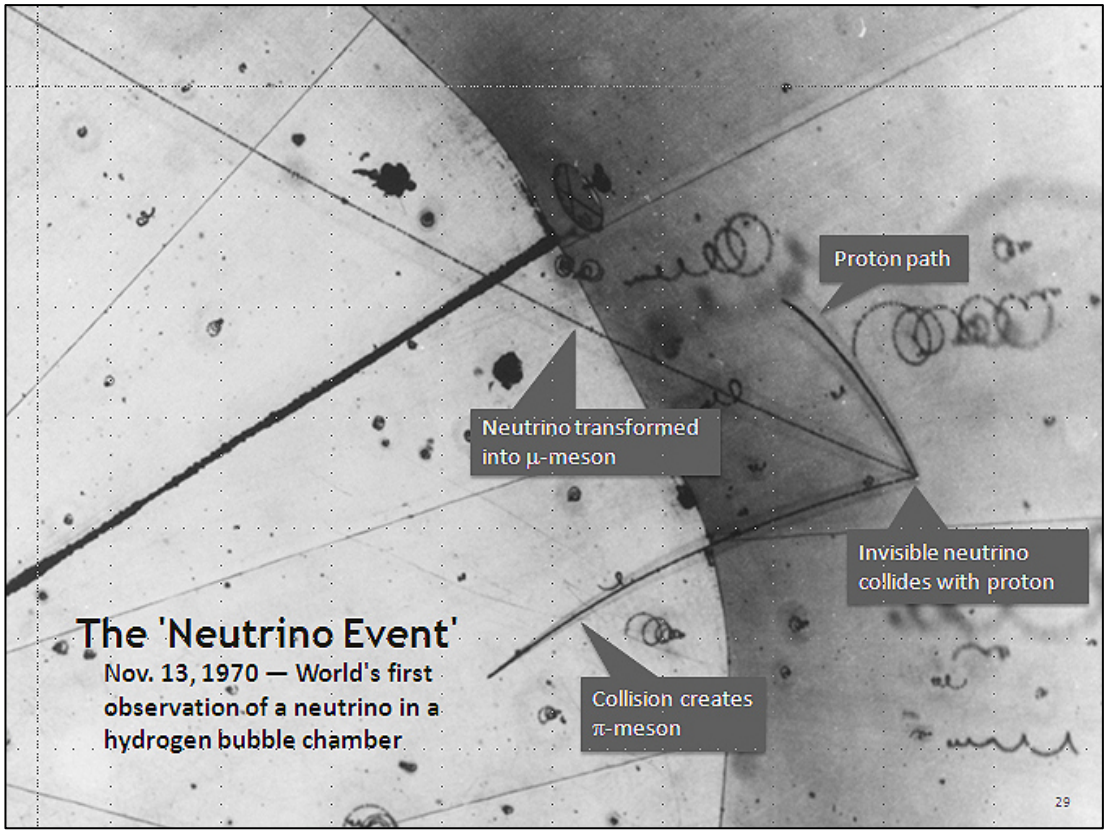
From Eric Torrence

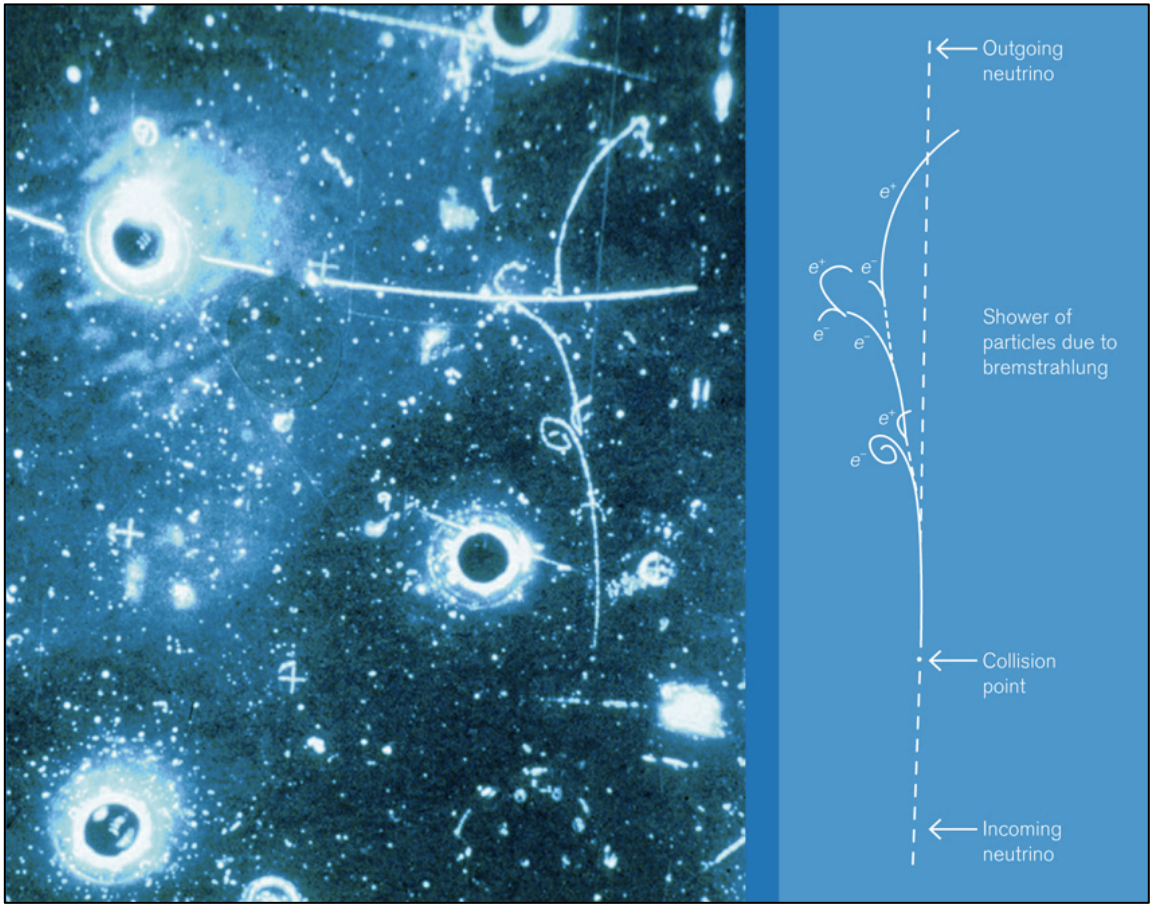


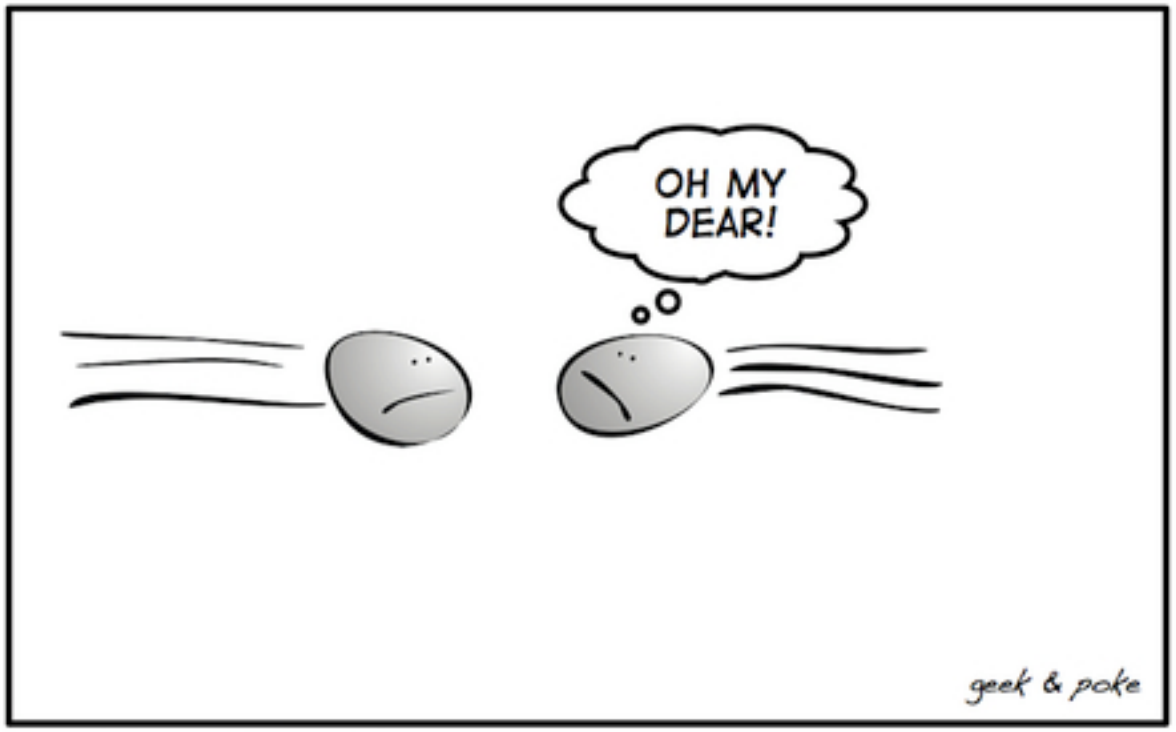
DATA QUALITY – DEFECTS





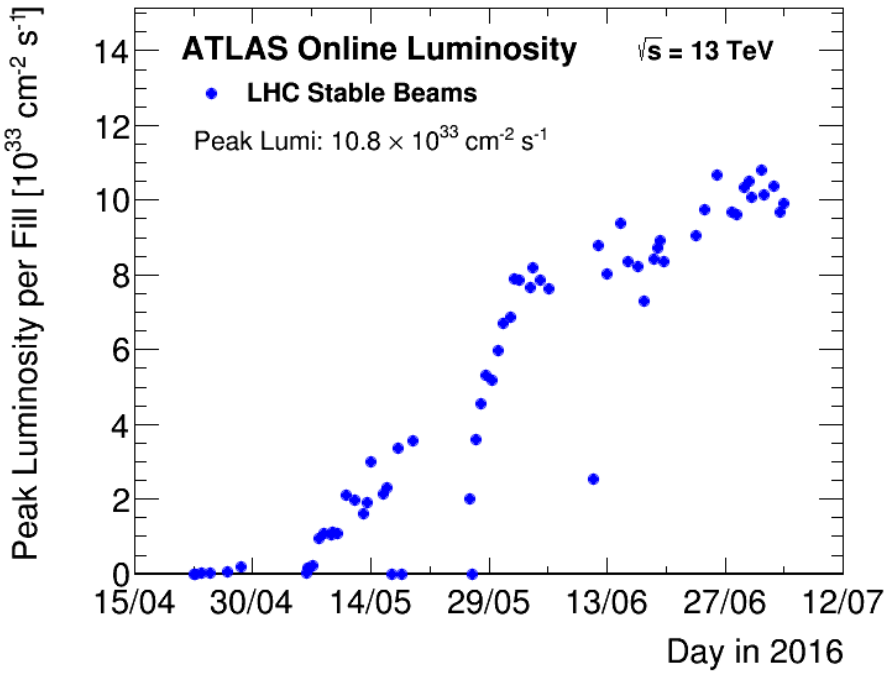

Proton path
Neutrino transformed into μ-meson
Invisible neutrino collides with proton
The 'Neutrino Event'
Nov. 13, 1970 — World's first observation of a neutrino in a hydrogen bubble chamber
Collision creates π-meson
29

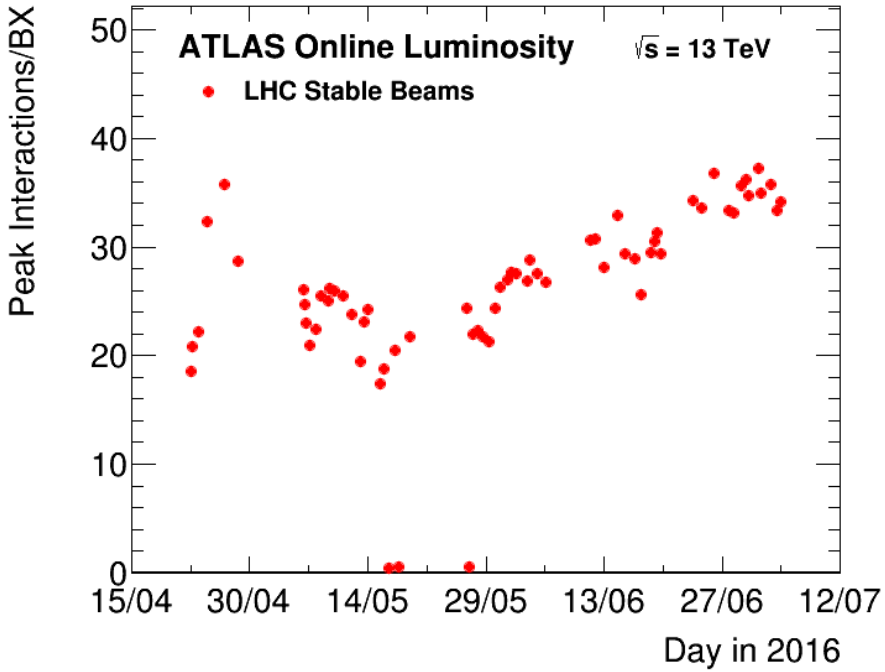

Outgoing neutrino
e^+
e^+
e^+
e^-
e^-
Shower of particles due to bremstrahlung
e^+
e^-
Collision point
Incoming neutrino



LATELY INSIDE THE LHC:
2 PROTONS 0.00000000000000000001 SEC BEFORE THE COLLISION

INT / XING


ATLAS Online Luminosity
√s = 13 TeV
LHC Stable Beams
Peak Lumi: 10.8 × 10^33 cm^-2 s^-1
Peak Luminosity per Fill [10^33 cm^-2 s^-1]
0
2
4
6
8
10
12
14
15/04
30/04
14/05
29/05
13/06
27/06
12/07
Day in 2016


ATLAS Online Luminosity
√s = 13 TeV
LHC Stable Beams
Peak Interactions/BX
0
10
20
30
40
50
15/04
30/04
14/05
29/05
13/06
27/06
12/07
Day in 2016

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.
- >> 50 combinations of MC generators in a sample.
- ~ 35 K samples generated on ATLAS in the last “campaign” of 2012.
- ~ 7 B events!

MONTE CARLO PRODUCTION CHAIN

How much processing time needed for each step?

Event Generation
simulate the physics process.

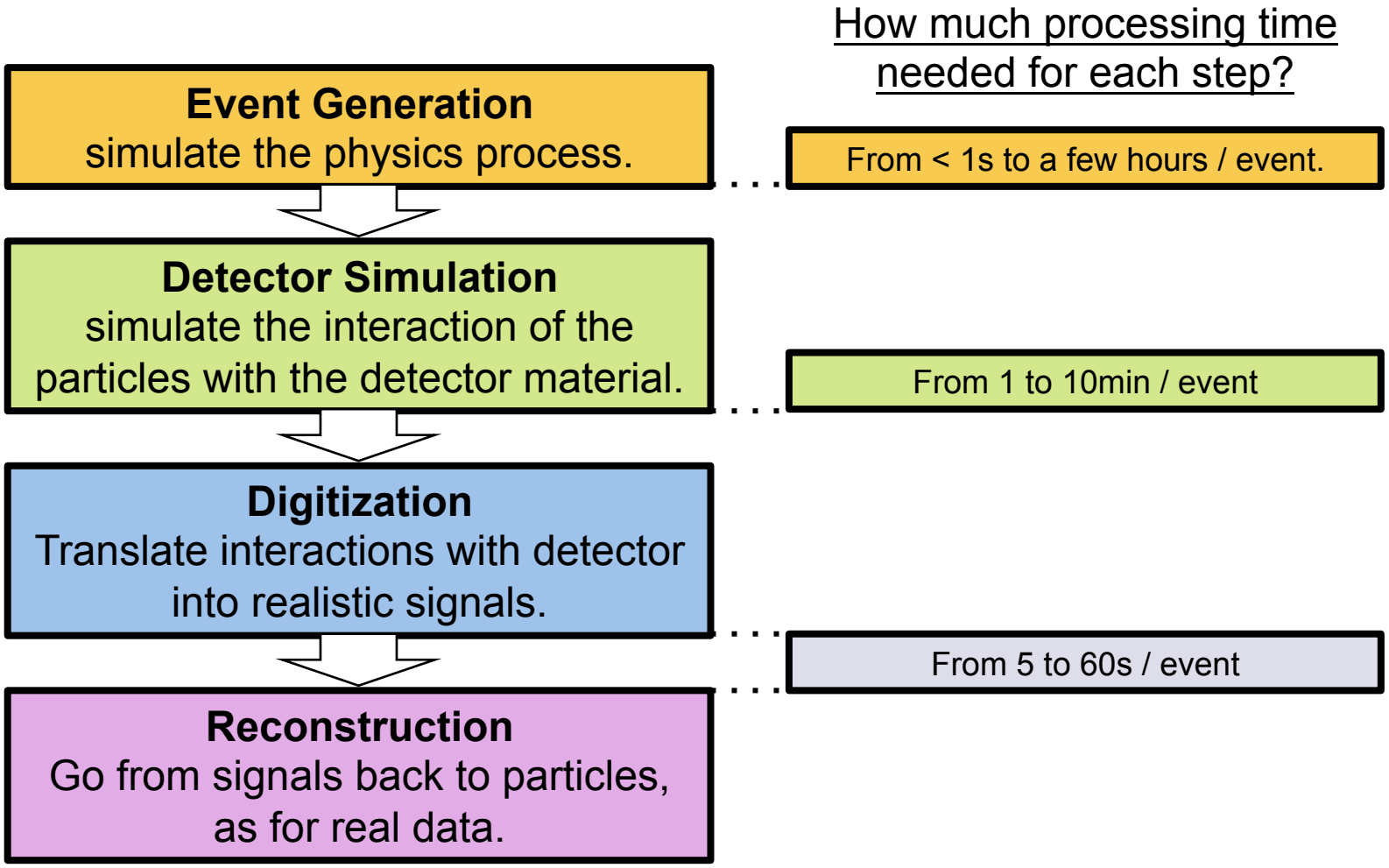
. . .

From < 1s to a few hours / event.

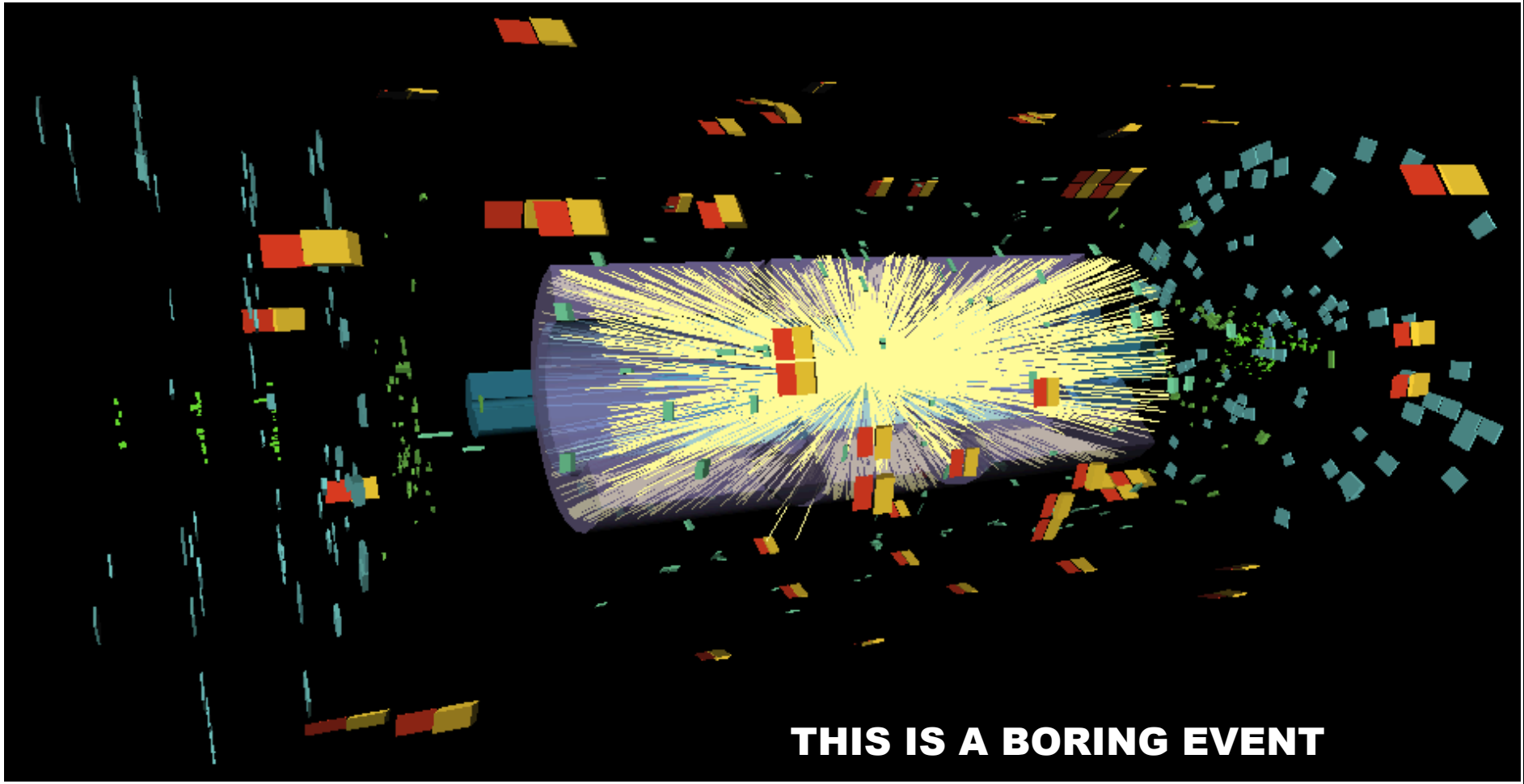
- ~ 50 MC generators on the market. *How many can you name?*
- >> 50 combinations of MC generators in a sample.
- ~ 35 K samples generated on ATLAS in the last "campaign" of 2012.
- ~ 7 B events!

Courtesy: Z. Marshall

MONTE CARLO PRODUCTION CHAIN



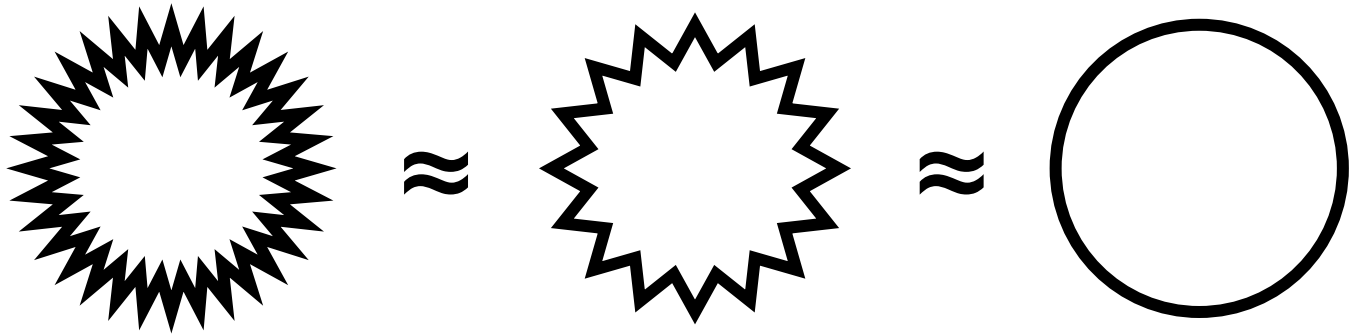
HOW TO SIMULATE THIS?



(Simulation material by Z. Marshal)

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.



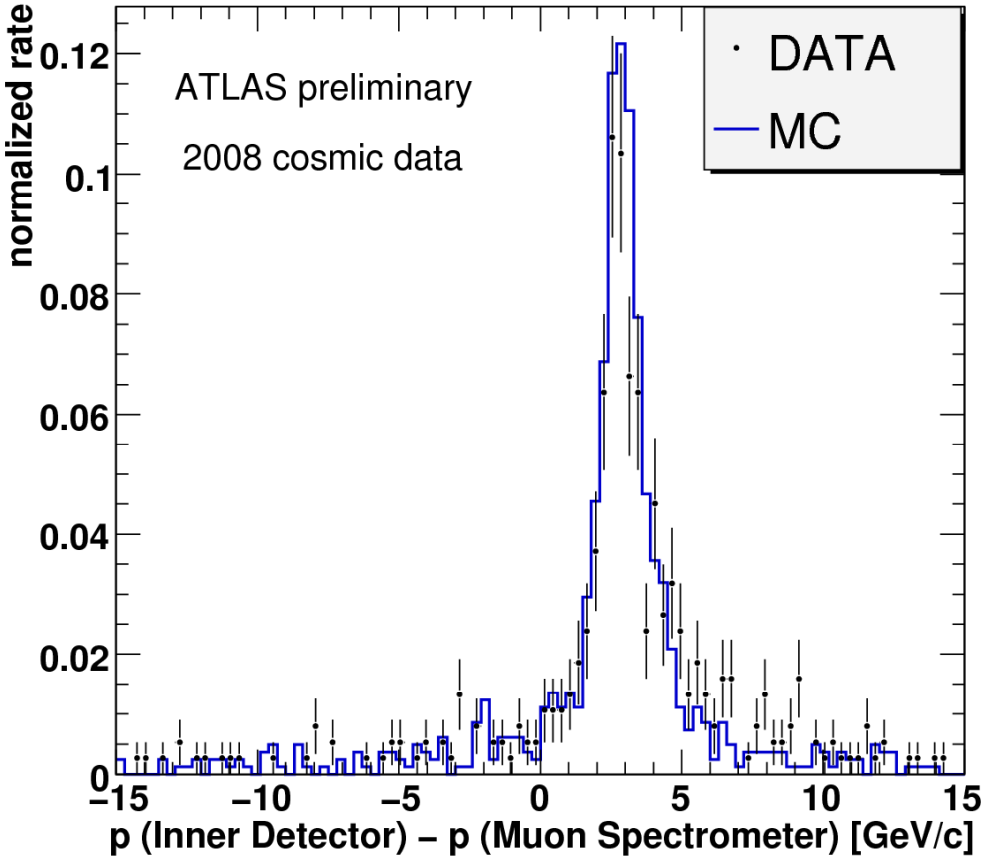
4. **Cross your fingers and press the button.**

HOW DO YOU KNOW IT WORKED?

- When the simulation can recreate something it *was not designed for*, you're doing well…

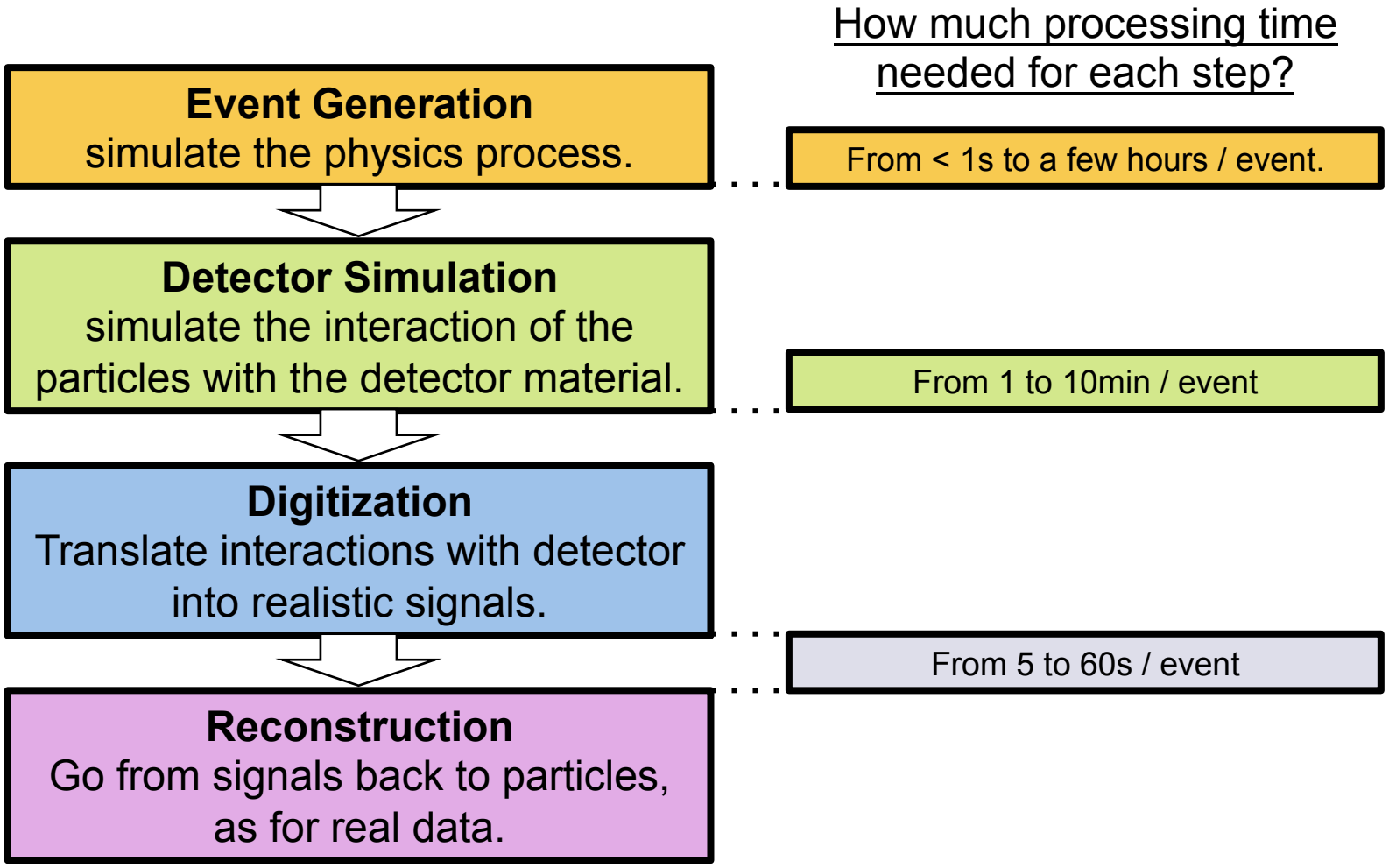
Cosmic rays are one interesting test. Use the simulation to propagate muons from the Earth's surface to the detector!

Here: energy loss in the calorimeter by a muon



(Simulation material by Z. Marshal)

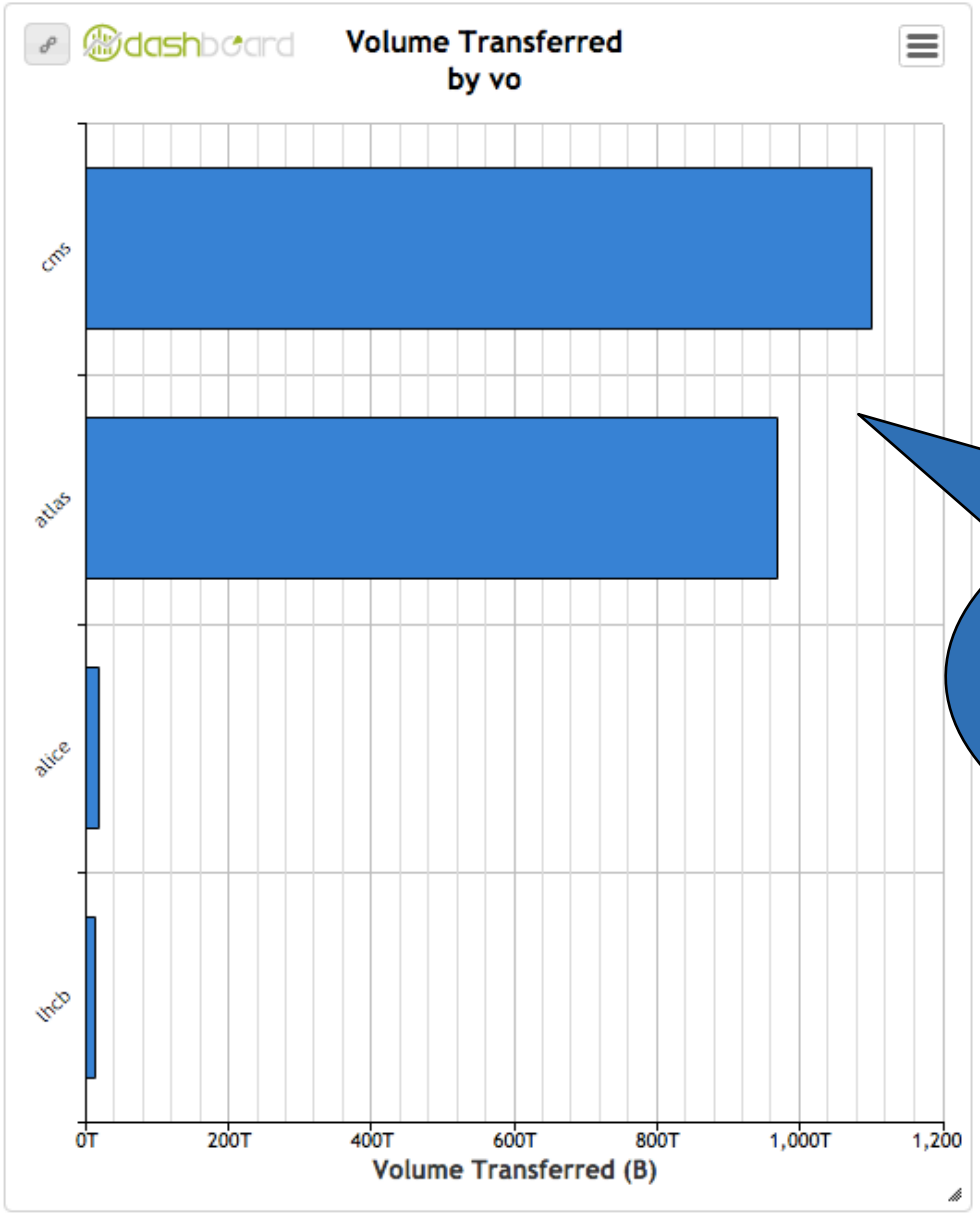
MONTE CARLO CHAIN



WLCG Transfers Dashboard

TRANSFER (2016-07-04 09:00 to 2016-07-05 09:00 UTC SLIDING) **PLOT:** METRIC | GROUPING | SERIES | ORDER | SIZE | STYLE

Matrix | Transfer Plots | Transfer Ranking | Expert Views

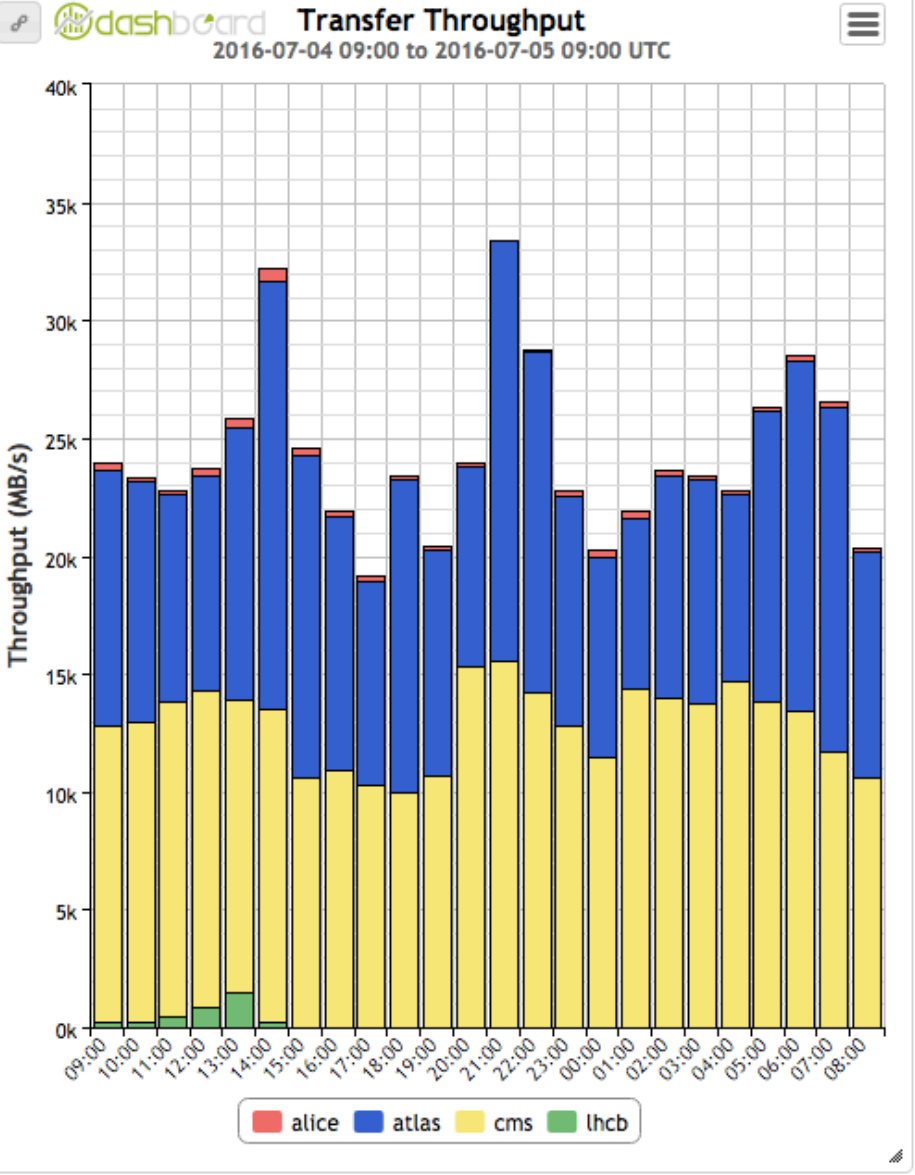


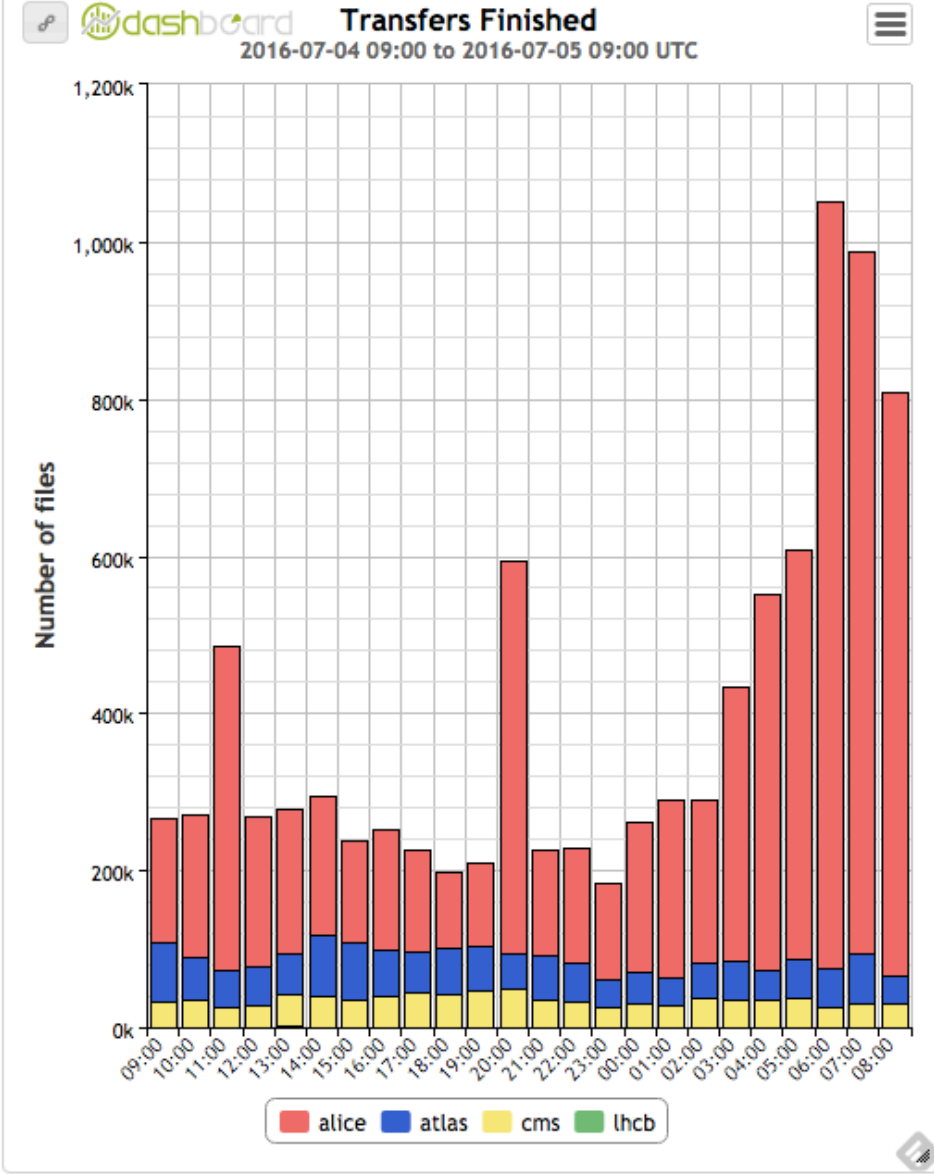
Big LHC experiments
transfer
O(1 PetaByte) / day!

WLCG Transfers Dashboard

TRANSFER PLOTS (2016-07-04 09:00 to 20 | GROUPING | TYPE | SERIES | SIZE | STYLE | BIN: SIZE | FORMAT | STEP

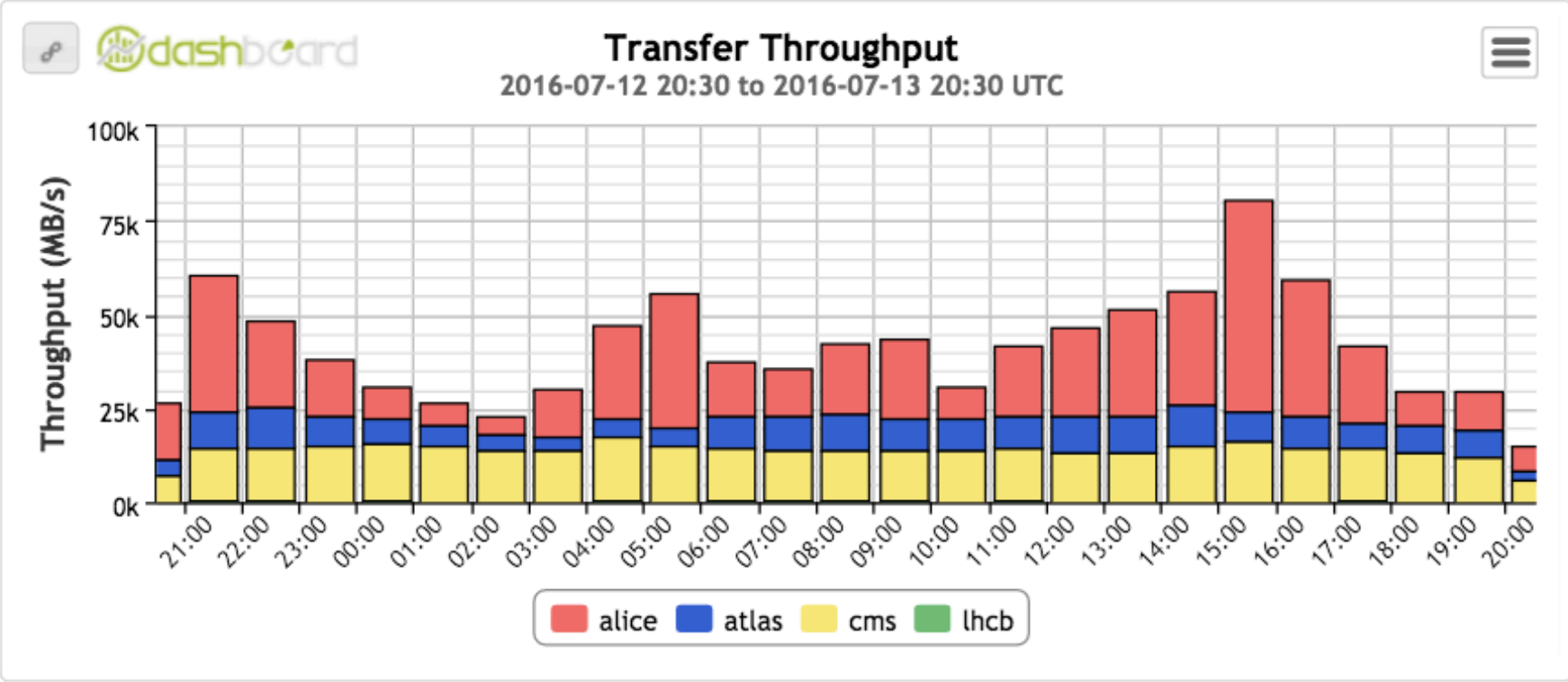
Matrix | Transfer Plots | Transfer Ranking | Expert Views





Welcome to the Worldwide LHC Computing Grid

Last 24 hours



The Worldwide LHC Computing Grid (**WLCG**) project is a global collaboration of more than 170 computing centres in 42 countries, linking up national and international grid infrastructures.

The mission of the WLCG project is to provide global computing resources to store, distribute and analyse the ~30 Petabytes (30 million Gigabytes) of data annually generated by the Large Hadron Collider (LHC) at CERN on the Franco-Swiss border.

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