Discussion on Run1 lessons

Physics at LHC and beyond

This is NOT a summary

This is NOT a prioritized list

This is just a collection of personal notes coming out of the talks or discussions during today.

Given the “online” feature of these notes please apologize any confusion/mistake.

Discussion start points

Starting from the major questions raised by the organizers

"Surprises and disappointments during run 1" (mostly performance-wise).
(i) What are the reasons for the "early" Scalar boson discovery wrt earlier expectations?
(ii) Operation modes (reprocessing, calibration, …).
(iii) MC usage (new features w.r.t. stability).
(iv) data/MC agreement. etc."

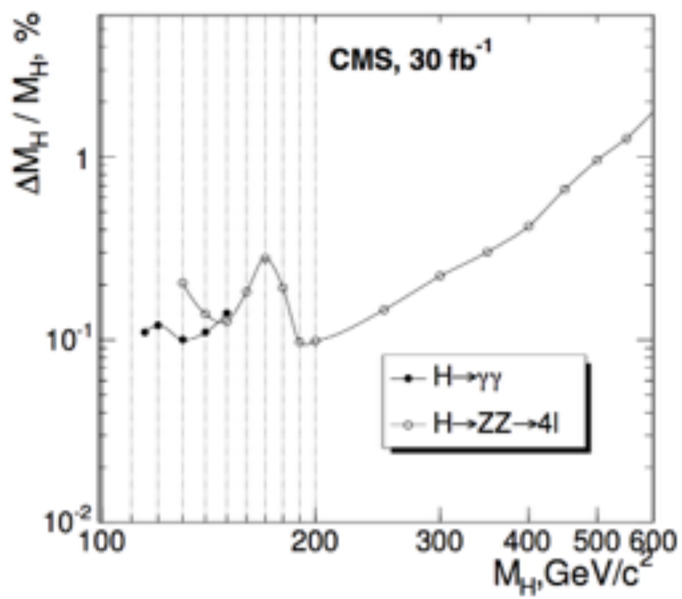
Not clear to me if we had real disappointments in Run 1 apart from Nature that did not give us much more than "a" Higgs (but this is not part of the performance discussion)

Trying to have a look at what we wrote back at the time of physics TDR:

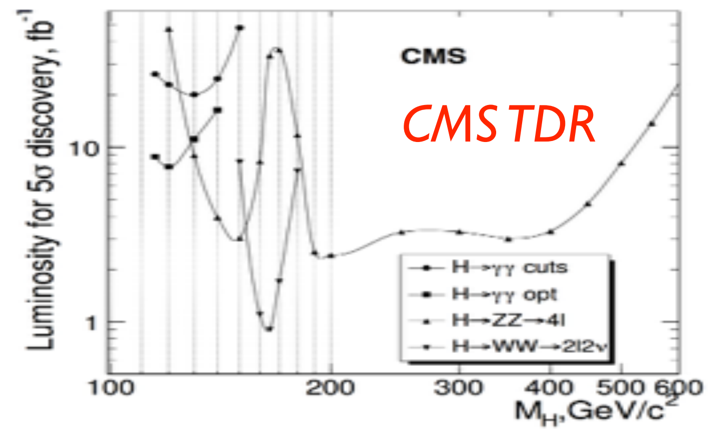
- all expectations (and in particular Higgs searches performance) were improved by a factor of ~2-4 (i.e. we have reached the same level expected with ~40fb^{-1})
- there are few exceptions (Higgs mass) were we exceeded the expectations by far
- there is at least one "bad" exception (W mass) were we are not even at the level we claimed possible with 10fb^{-1}
  - ➡assumed different PU?
  - ➡efforts just up to the level of the biggest systematics for Higgs? Momentum scale, etc.

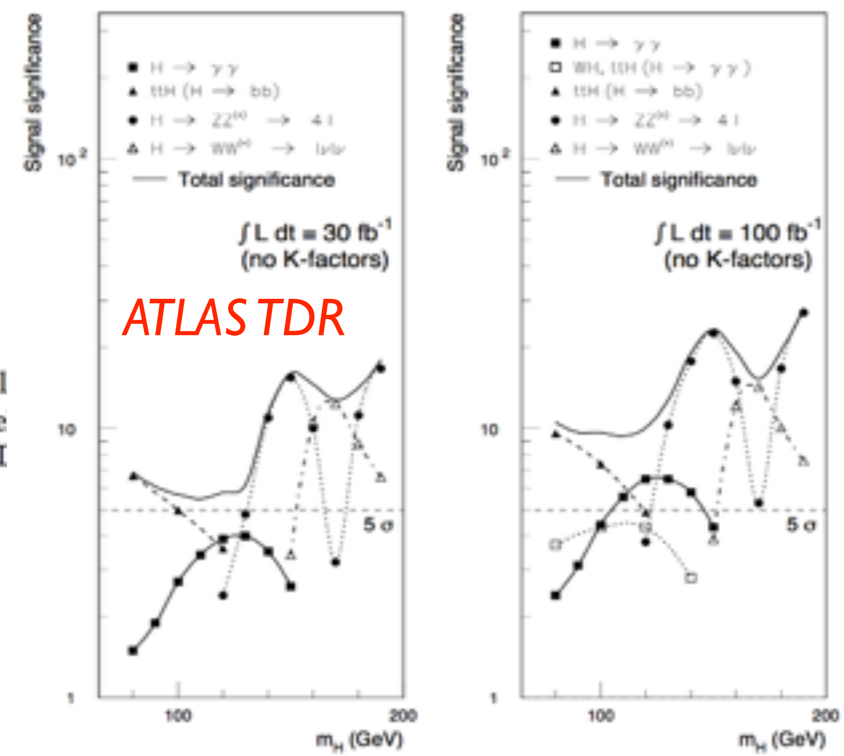
Successes

From physics TDR we were expecting to discover the Higgs with factors more the statistics we have used and in conditions that were much “easier” than what we had:



All the studies reported in this TDR include the effects of pile-up on the signal. For al with luminosities up to 60 fb^{-1} $\mu = 5$ was used. Several techniques have been deve minimise the effect of pile-up, and have been used in the studies reported in this TI in-time and out-of-time pile-up has been included.





The exception

Source of uncertainty	uncertainty with 1 fb^{-1}	ΔM_W [MeV/c^2] with 1 fb^{-1}	uncertainty with 10 fb^{-1}	ΔM_W [MeV/c^2] with 10 fb^{-1}
		scaled lepton-p_T method applied to W$\rightarrow e\nu$		
statistics		**40**		15
background	10%	10	2%	2
electron energy scale	0.25%	10	0.05%	2
scale linearity	0.00006 / GeV	30	<0.00002 / GeV	<10
energy resolution	8%	5	3%	2
MET scale	2%	15	<1.5%	<10
MET resolution	5%	9	<2.5%	< 5
recoil system	2%	15	<1.5%	<10
total instrumental		**40**		**<20**
PDF uncertainties		20		<10
Γ_W		15		<15
p_T^W		30		30 (or NNLO)
		transformation method applied to W$\rightarrow \mu\nu$		
statistics		**40**		15
background	10%	4	2%	negligible
momentum scale	0.1%	14	<0.1%	<10
$1/p^T$ resolution	10%	30	<3%	<10
acceptance definition	η-resol.	19	$< \sigma_\eta$	<10
calorimeter E_T^{miss}, scale	2%	38	≤1%	<20
calorimeter E_T^{miss}, resolution	5%	30	<3%	<18
detector alignment		12	–	negligible
total instrumental		**64**		**<30**
PDF uncertainties		≈20		<10
Γ_W		10		< 10

Most likely this was a victim of two killers:

- the prejudice that at pp colliders this is an impossible measurements (i.e. dominated by theory syst.)
- the Higgs search: everything was studied and improved manly for the Higgs (i.e. no need to have a momentum scale on the muons below 10 MeV)

Performance comparison with expectations

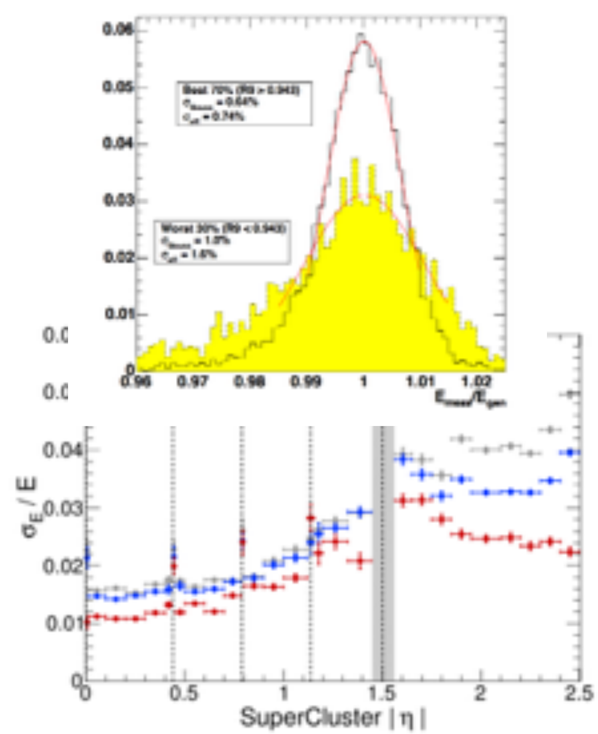
ESSENTIAL to focus on "particle" performance more than proxies (as jet, MET, Energy deposits, tracks, etc):

- particle ID is at the core of LHCb physics program
- Particle Flow in CMS allowed to exceed initial expectations but similar approaches in other experiments did not show up as a big jump (need an appropriate design!)
  - big magnetic field to open up tracks
  - excellent tracking capabilities up to very low pT (how low pT? See Daniel's question)
  - calorimeter resolution and granularity just enough to make proper links

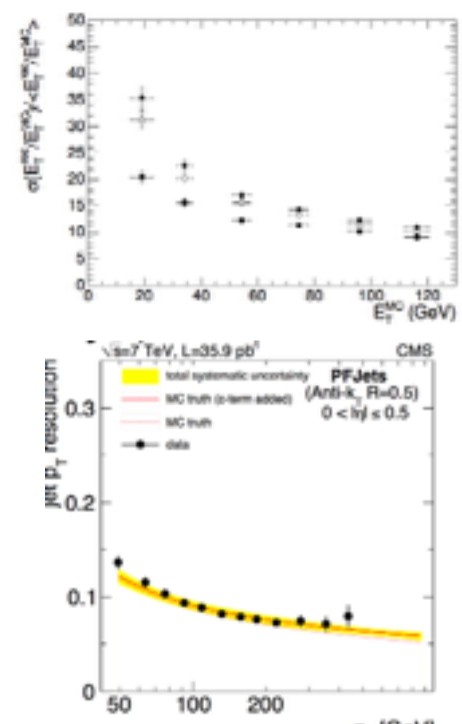
A side note:

we did much better in physics result w.r.t. pTDR expectations especially in cases where the "object" performance was improved

CMS exp/results photon



CMS exp/results jets



Potential list of topics to discuss

We have performed (much) better than foreseen and much faster w.r.t. past experiments. What can be the reasons?

- Proper tools/techniques:
  - particle flow as main reconstruction paradigm (see before)
    - ➡ why ATLAS and CMS differ? See Lindsey presentation
    - ➡ ATLAS also performs some sort of PFlow for MET (linking tracks to unassociated clusters… see Tai's presentation)
    - ➡ similar performance in Jet/MeT, what about lepton isolation and PU subtraction?
    - ➡ is low pt tracking strictly needed everywhere or we can be more creative and differentiate tracking depending on situations so to optimize (see CMS jet core tracking iteraton) - see Daniel's question
  - simulation (starting from GEANT)
  - software optimisation
  - speedy reprocessing

Potential list of topics to discuss

We have performed (much) better than foreseen and much faster w.r.t. past experiments. What can be the reasons?

- Operation/preparation: calibration/alignment well thought and workflows immediately available
  ➡ to avoid any comparison between current detectors, in L3 it took three years to understand how the tracking detector worked (probably not a good example, but a quite painful personal one…)
  - Open discussion on other more creative ways
    - data scouting?
    - different processings according to selected triggers
    - abandon the paradigm of "raw data" and make analysis "online" ?

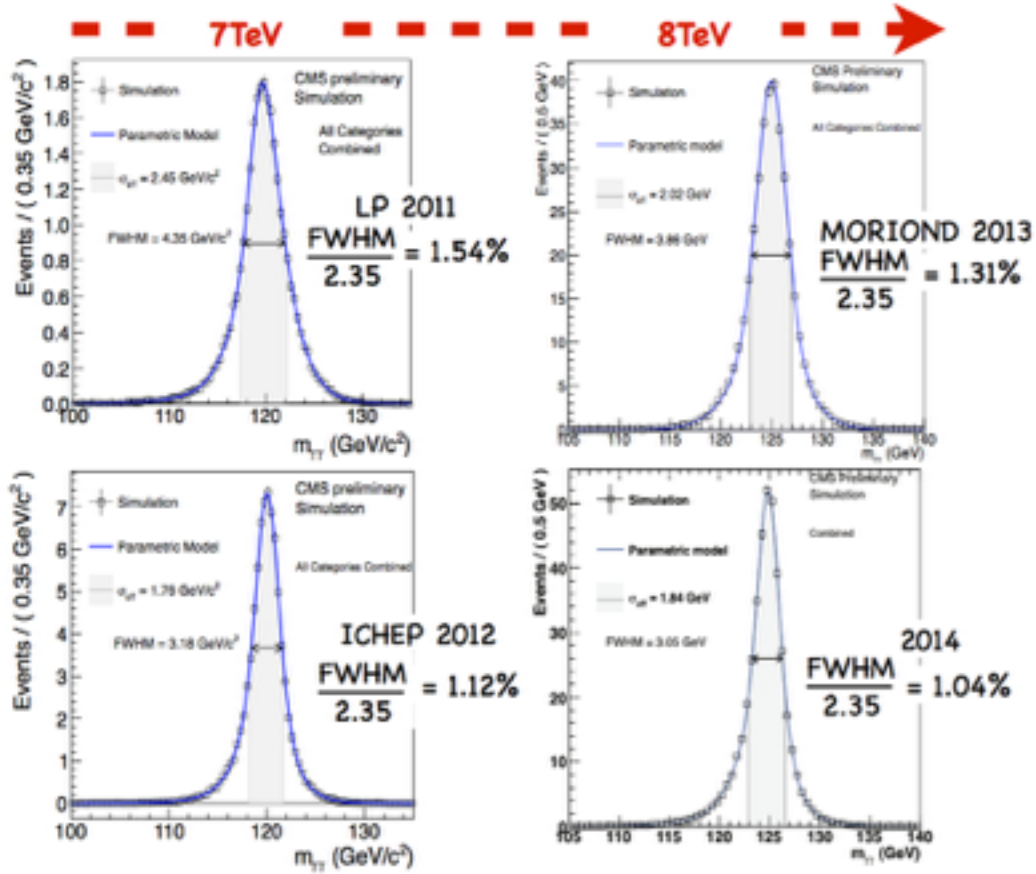
- MVA techniques (this was a paradigm change: would have the Higgs been discovered so early without ?)
  - see for example btag in CMS, why a less efficient tagger is used?
- the early incident (that gave us one more year to study) ?
- what else ?….

Lessons learned: resolution

- **Long journey to improve the energy resolution:**
  - ✓ **improved calibration of the ECAL detector**
  - ✓ **improved description of the ECAL simulation with a run-dependent Monte Carlo description of the detector that follows the evolving conditions during data taking in 2012, and includes the simulation of out of time pileup over the time windows [-300 ns, +50 ns]**
  - ✓ **improve multivariate energy correction using a semi-parametric likelihood technique in order to construct a prediction for the full distribution of E-True/E-Raw.**

PROMPT reconstruction within 48h from data taking

RECONSTRUCTION with improved conditions



Daniele Benedetti (Purdue)

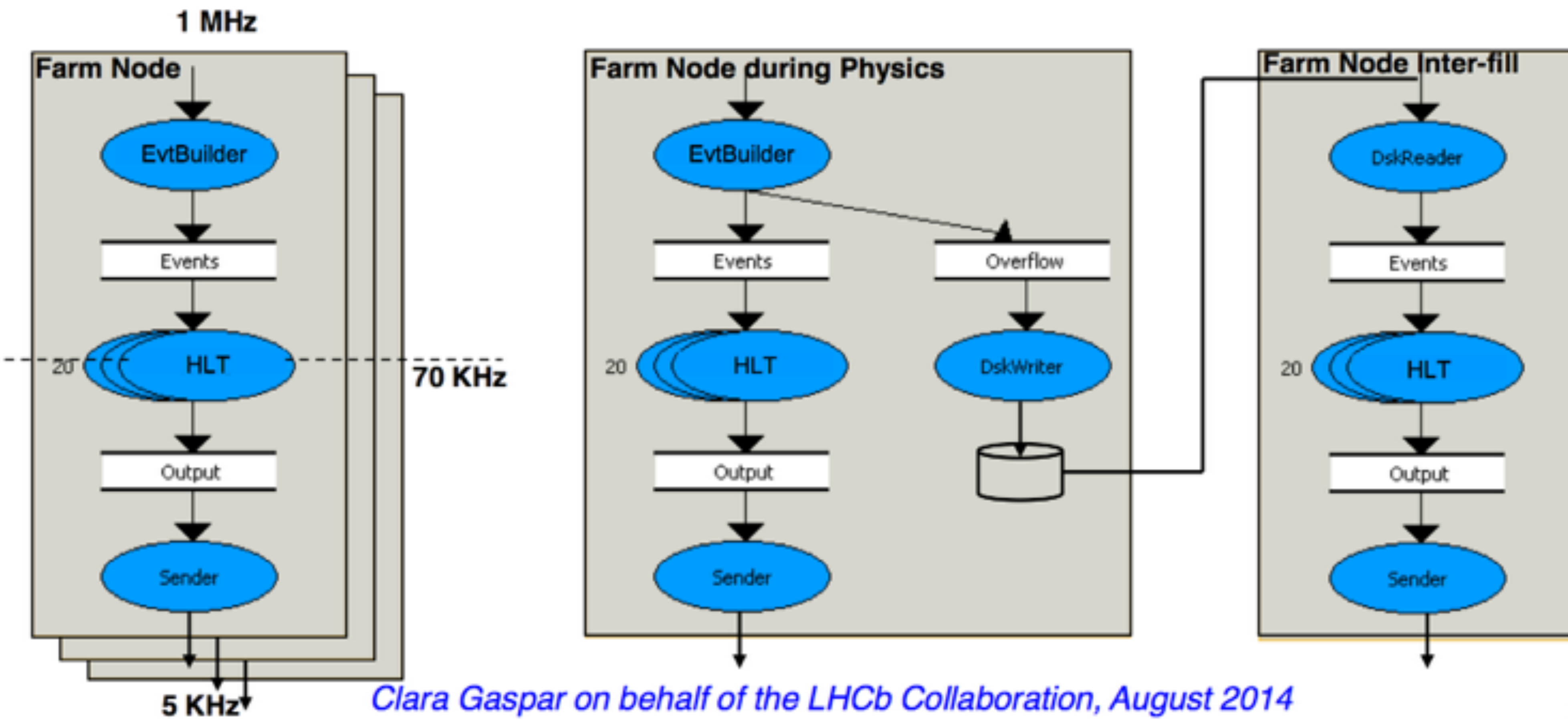
16/17

HLT Farm Usage

- **Resource Optimization -> Deferred HLT**
  - Idea: Buffer data to disk when HLT busy / Process in inter-fill gap
  - Change of "paradigm": A run can "last" days…

Standard HLT

Deferred HLT



Clara Gaspar on behalf of the LHCb Collaboration, August 2014 11

Trigger

The red or the blue pill

more comprehensive and generic single object (necessarily high threshold)

Pro:
- simple to implement and maintain
- easy to combine at analysis level
- suffering less systematics effects
- easy to simulate

Con:
- need to keep thresholds high
- high rate and high final statistics sample to start with
- difficult to cover special use cases (compressed spectra, complex signatures)

specific and analysis optimized: bring algorithm to HLT and possibly L1

Pro:
- difficult to implement and maintain
- difficult to combine at analysis level
- might be affected by syst.
- difficult to simulate

Con:
- allow to lower thresholds
- manageable rates and finla statistics sample
- aimed to cover special use cases (compressed spectra, complex signatures)

Alternative approaches? Data scouting, online analysis, dataset specific reco,…

Pile-up

Several techniques used and in preparation:

- statistical subtraction
- reweighing techniques
- vertex-aware techniques (charged particle based subtraction)

Is this a reason in favor of "tracking" calorimeter for the future?
(see ATLAS/CMS Hgg analysis comparison for example)

Future running conditions
- very little seen so far (and here) on out of time PU, but 25ns running is approaching
- how to implement a "time"-aware PU mitigation?

If time left…..

Was the design of the detector optimal?

"A posteriori" this is always an easy question to answer and from a point of view of the results the answer can only be "yes".

But if we look at the way the results were obtained there are few questions that arise and I would like to open for discussion
(please take it as a provocation):

- how simplicity in detector design improves the turn-around (data taking, calibration, re-re-re-processing, analysis)?
  - ➡BTW, it took two years to have the latest greatest calibration from ATLAS and CMS
- Is "better resolution" in the calorimeters an absolute paradigm or other things play an important role?
  - ➡a good tracking capability (and magnet) seems to compensate pretty well
  - ➡are granularity and longitudinal segmentation helping?
  - ➡vertexing capability? PU?
- What is the cost/benefits of having an high eta region as performant as the barrel?
  - ➡How many analyses really benefited, apart from hermeticity arguments (missing et, etc.)?