

2009 October Exercise: Study of ZZ->4l signal and backgrounds

* Andrius Juodagalvis (VU ITPA, Lithuania [↗](#), email)

Goals and Motivation

The October exercise was intended to check the readiness of the collaboration to work with real data. Another goal was to train people to do all the tasks pertinent to their analysis. My participation in the exercise was relaxed, since during the OctoberX period I wanted to process a larger number of Monte Carlo events related to my analysis (i.e. to increase the statistics). At the same time, doing the analysis in parallel to more-closely tracked activities set a background to evaluate my experience (with crab jobs in particular).

Previous presentation of the analysis: * EWK Multiboson meeting
08-October-2009 [↗](#)

The datasets in focus (H->ZZ->4l background datasets):

ZZ->4l (signal)	/ ZZ_4l_10TeV_GEN/ndefilip-CMSSW_2_2_7-10TeV_RAW/DIGI_RECOSI/M_I_DEAL
ttbar->4l (background)	/ TT_4l_10TeV_GEN/ndefilip-CMSSW_2_2_7-10TeV_RAW/DIGI_RECOSI/M_I_DEAL
Zbb (background)	/ LLBB_4l_10TeV_GEN/ndefilip-CMSSW_2_2_7-10TeV_RAW/DIGI_RECOSI/M_I_DEAL

Personal analysis code was used. The code requires CMSSW_2_2_13 full framework.

Workflow

The planned steps were:

1. Process datasets from RECO to PAT level requiring at least 4 leptons (electrons or muons, excluding cases 3+1) in the selectedLayer1. This step was applied on ZZ4l, ttbar and Zbb datasets. The ZZ4l dataset was additionally separated into signal and background samples based on genParticles information.
2. Publish the datasets in cms_dbs_ph_analysis_01 for the reuse.
3. Use the published datasets to reconstruct ZZ->4l candidates, applying cuts and constraining the invariant mass of any two leptons.
 - ◆ electron preselection: `cms.string("(electronID('eidRobustLoose') > 0) & (pt > 5.)")`
 - ◆ muon preselection: `cms.string("(isGlobalMuon = 1) & (((pt > 5.) & (abs(eta) <= 1.1)) | ((pt > 3.) & (p > 9.) & (abs(eta) > 1.1)))")`
 - ◆ *Note*: Since the code was developed to study selection efficiencies, less strict cuts on the cloned layers were also applied.
4. Collect the numbers and make some plots.

Experience during the analysis

The main focus was to run crab jobs at T2_FR_IN2P3 to increase the statistics from ~150k events to ~1M events for each data sample, including the processed data publication in the local DBS. An additional goal was to check the workflow. Earlier results were obtained on a local workstation.

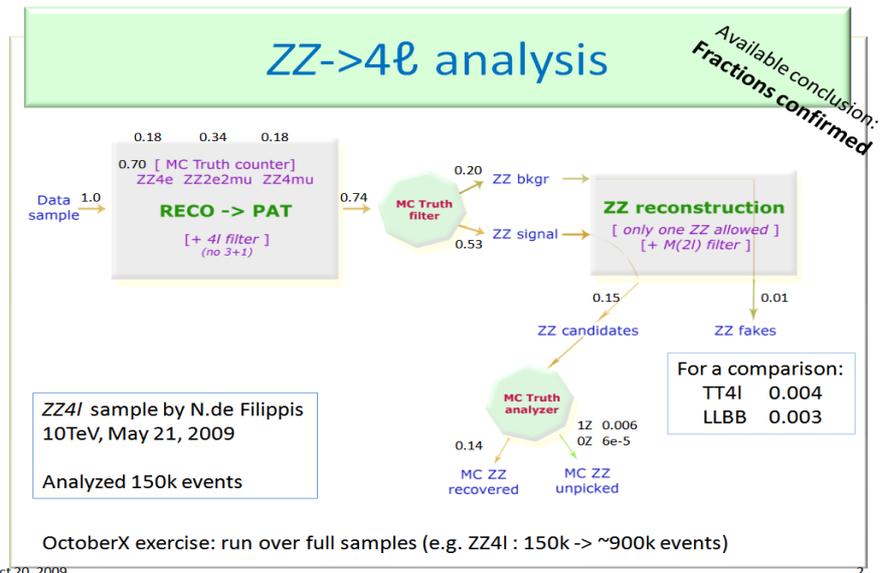
Since the code needed some development before the exercise, the analysis started in the evening October 8. The beginning was promising with many crab jobs completed by the afternoon of next day. A few jobs had an error 60303 (file exists) or the error was not specified. These jobs were resubmitted having removed the result file from the storage element. On Monday, the 12th, 3 processed sets (ZZ4l signal, ttbar and Zbb) were already published, 1 ZZ4l background job was holding from completion of Step 2. In addition, the site T2_FR_IN2P3 was (temporarily) not in the production mode. On Tuesday, the 13th, the last job from Step 1 was completed, Step 2 was passed, Step 3 jobs had 100% failure (jobs terminated while in a queue). Thanks to the efforts from the support personnel, a site problem with improperly mapped "ordinary users" was identified and fixed. Crab 2.6.3 patch 2 was also released, so Step 3 jobs ran on Wednesday with 100% of success. Making plots required additional cmsRun jobs, thus the processed datasets were copied (using lcg-cp) from the SE to the local workstation, which turned out to be not very smart choice having in mind the time constraint of the exercise and personal schedule. The difficulties resulted from several sources. Some files were not copied with the first attempt and their absence was not noticed until several attempts to process the data were done. Several attempts to run cmsRun jobs were needed due to the failures on a local workstation. In addition, a single-CPU was rather slow to process a sheer number of events.

Some summarizing remarks:

1. The final plots were not produced lacking a definite goal what would constitute a new result in comparison to other studies. This also discouraged from a continued use of the personal code when similar codes are already developed.
2. Difficulties related to crab jobs were mainly the same as experienced by other participants of the OctoberX exercise (completed crab jobs still in 'submitting' status, code 60303, and the like).

Results

The number of events passing through different steps of analysis are



shown in the plot.

-- AndriusJuodagalvis - 26- Oct - 2009

This topic: Main > AndriusJ2009OctoberX

Topic revision: r1 - 2009-10-26 - AndriusJuodagalvis



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback