Summer Student Programme 2009 · DESY

Project Report

The Online Event Display for CMS Centre at DESY

Daniele Zanzi, University of Bologna*

Supervisor: Günter Eckerlin, DESY

September 7, 2009

Abstract

The main goal of the project is to provide the CMS remote centre at DESY with an online event display. This is a crucial tool that can enable the DQM shifters to evaluate the performances of the detector and its components in real time and to check the quality of the events which are being taken. Another part of the project consists in the creation of web-based calendars in order to organize the activities of the remote centre.

*danielezanzi@gmail.com

Contents

1 The CMS Experiment

The Compact Muon Solenoid (CMS) is one of the four experiments installed along the ring of the Large Hadron Collider (LHC) [1] at CERN. It is one of the biggest and multi-purpose detectors like its cross-check ATLAS.

The aim of this machine is to unveil the missing building blocks of the Standard Model. This theory, indeed, is one of the most powerful and well tested among the physical theories, but till now it can just be considered as an effective model. This general framework is based on lots of parameters, like the masses of the fermions, which can't be explained from a fundamental level, and on some mechanisms, like the spontaneous breaking of the electroweak symmetry, which implies the existence of particles not yet observed.

The research of one of these missing particles among others, the Higgs boson, inspired the development of LHC and CMS.

1.1 The Physics

In order to explore the mass scale in which the Higgs boson is likely to be found, in the LHC tunnel two proton beams will collide. This choice allows to ramp the energy up to 7 TeV for each beam, but since protons are not fundamental particles the events are not clean as in the e^+e^- collision and a lot of hadronic background is produced.

Due to this scenario, not all the possible decays of the Higgs boson can be studied and since the branching ratios (fig. 1) are strongly dependent on its mass, m_H, CMS should be able to detect several decay modes with completely different products.

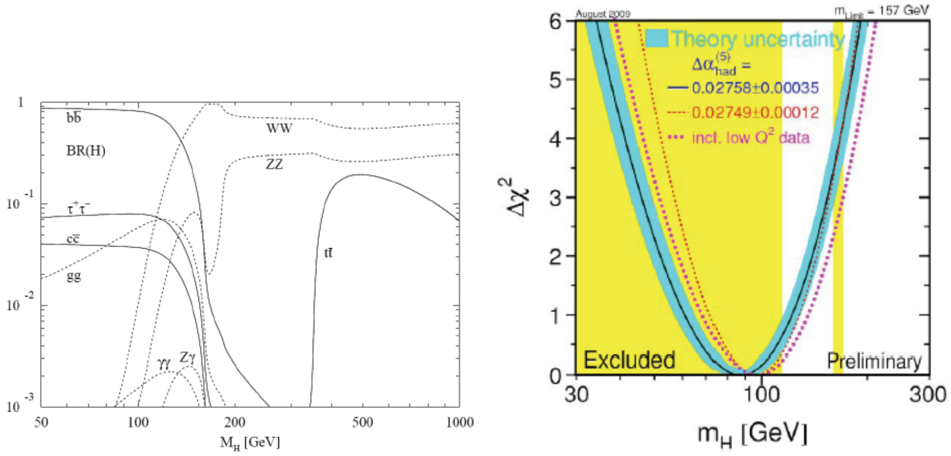

Figure 1: Branching ratios of Higgs boson for different masses, on the left, and the χ^2 for m_H from the fitting on the SM data with the excluded mass ranges in yellow, on the right.

There are three main mass ranges:

1. *Low mass region* (114.4 GeV/c^2, LEP limit $< m_H <$ 130 GeV/c^2): the dominant decay is $b\bar{b}$, but the QCD background of the LHC events make this signal completely hidden. One of the most clean decay mode is $H \rightarrow \gamma\gamma$ even if the branching ratio is very small.

2. *Intermediate mass region* (130 GeV/c^2 $< m_H <$ 600 GeV/c^2): under the threshold of $2m_Z$, the best channels are $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow 2l2\nu$[1]. In both cases the final products are four leptons and this kind of signal has little background. CMS, in particular, is designed to perform very well with muons. Above $2m_Z$ these two channels become on-shell and the $H \rightarrow ZZ \rightarrow 4\mu$ is the most promising.

3. *High mass region* ($m_H >$ 600 GeV/c^2): the cross section decreases and so it is necessary to take into account also channels in which Ws or Zs decay in jets.

In addition to this leading research, CMS will also look for new physics trying to solve important enigmas like the Dark Matter, the Dark Energy and the SUperSYmmetric particles.

1.2 The Detector

In order to fulfill all the physical requirements, CMS has to be a multi-purpose detector with the characteristic onion-like structure focused mainly on an high performance electromagnetic calorimeter and a good muon resolution. The demand on the muon reconstruction forces the detector to have a high magnetic field and

[1]The asterisk means that one of the two bosons is virtual.

that mean that the volume inside the magnet can't be very big. Consequently, all the sub-detectors have to exploit as better as possible the available space.

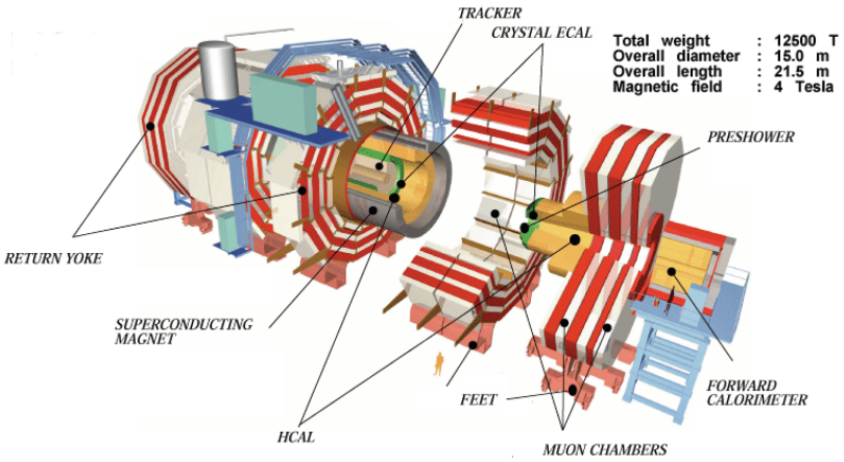

Figure 2: CMS detector structure.

Starting from the inner part, the sub-detectors which compose CMS are listed:

- **Inner Tracking System**: the two main issues are the high track multiplicities and the high rate of radiation. These require high granularity and radiation hard systems. Consequently, for vertexing there are 3 layers of silicon pixels, while for the tracking there are 10 layers of cooled silicon microstrip detectors.
- **EM Calorimeter**: The lead tungstate ($PbWO_4$) is the scintillating material, thanks to the short radiation and Moliere lengths, to the fast light emission and to the radiation hardness. The drawback is the low light yield. So the light has to be amplified by silicon avalanche photodiodes in the barrel region and by vacuum phototriodes in the endcaps. The calorimeter is also supplied with a preshower system for π^0 rejection.
- **Hadron Calorimeter**: since the central calorimeter is inside the magnet coil, brass has been chosen for the absorbing material due to the short interaction length and its non-magnetic property. To maximize the interaction lengths the active material is a plastic scintillator tile read out by wavelength-shifting fibers. In order to get a full geometric coverage for the transverse energy there are a forward calorimeters made of iron and quartz-fibers.
- **Muon System** The return magnetic field is large enough to fill 1.5 m of iron around the magnet. In this volume 4 muon "stations" are installed. Each of them is composed of aluminium drift tubes in the barrel and cathode strip chambers in the endcaps. In both the regions there are also resistive plate chambers in order to improve the time resolution and to identify the bunch crossing.

1.3 Experimental Challenges

The event rate at LHC is one of the most striking challenge. At center-of-mass energy of $\sqrt{s} = 14$ TeV, the total *pp* inelastic cross section can be extrapolated around 70 mb. At the design luminosity of 10^{34} cm^{-2}s^{-1} the expected rate is about 10^9 events per second. Moreover the rate of the bunch crossing is 40 MHz, which means that each 25 ns two bunches collide and in each collision about 20 events are piled up.

This scenario is a real challenge for physicist for three main reasons. First, the BC rate is so high that the particles are still flying through the detector when the next collision occurs. Second, the piling up of different events requires an high granularity in the detectors. Third, due to the high number of channels the amount of data from the sub-detectors is so huge that it can't be completely stored.

For these reasons a very sophisticated two-level trigger reduces the rate of events which are stored for the analysis from 10^9 to 10^2 per second, while a very good time resolution is needed to synchronize the millions of channels and to not confuse signals from different bunch crossing.

2 Data Quality Monitoring

The challenging features of CMS need of an effective Data Acquisition system (DAQ), i.e. a system that collects all the data read out from each channel of the detector and builds up the events.

This system has to be highly reliable and robust because it plays a crucial role. Through the two levels of trigger, the DAQ system has to select and build up the interesting events at a rate of 100 Hz with a rejection power of 10^7.

This means that the selection has to be always kept under control. And this task is undertaken by the Data Quality Monitoring system (DQM) [2]. Its goal is to provide a common infrastructure usable by different groups in the experiment for checking the quality of the data taking. DQM applications can be used on-line in the High Level Trigger farm, or in a local DAQ supervision of a particular sub-detector, or even off-line, for instance, in the "production validation".

DQM on-line system collects information from the HLT algorithms and processes them independently from the HLT farm. In this way the data acquisition process can't be affected by eventual DQM malfunctions.

The processed informations are sets of objects, called "monitoring elements" (ME), for instance histograms, scalars, string messages and so on, and the data flow is organized in three steps: the sources, or "monitoring producers" which create the MEs; the "collectors" which transfer data and provide communications between the other two levels; the clients, or "monitoring information consumers", which are the final users of the architecture.

Since the main goals of the DQM applications are to check the quality of data in the read out stream from the detector to the storage elements, the system doesn't provide the possibility to handle specific events, but only monitoring elements. In this way the monitoring activity is completely separated from the DAQ processes by the collectors.

All the MEs contains informations about statistical samples of events. Monitoring elements from the same DQM application can be different, but, as mentioned afore, permit to evaluate the activity, for instance, of particular sub-detectors.

The DQM system provides not only on-line but also off-line applications. The off-line DQM activity is mainly involved in the "production validation". All the data taken during the CMS runs are reprocessed within the next 24 hours and validated. The algorithms are mostly fast reconstruction methods similar to that ones used for the HLT. The aim of this procedure is to check the consistency of the data in order to store and let them be available for physics analysis.

2.1 DQM Control Room

There are three control rooms for the on-line DQM analysis: one at Meyrin (P5), next to the experiment, one at Fermilab and one at DESY. The most important is the control room in P5, where the DQM shifter sits next to detector experts. The direct feedback between these people is crucial, especially in the initial part of the life of the experiment, because on one side it provides people responsible of sub-detectors with realtime informations about the performances of the hardwares, on the other side let the DMQ person to interpret more easily the monitoring elements.

After the initial period, the other two remote control rooms will become more important in order to have the 24-hours full coverage of the run time and they will carry on autonomous activities.

3 The DQM Event Display

One of the most effective way to analyze the overall behavior of detectors and sub-detectors is through an event display. The visualization is fundamental for all the kind of data analysis and the CMS collaboration developed a very flexible Interactive Graphics for User Analysis system (IGUANA) [5]. It is a generic HEP visualization toolkit and provides user with a Graphical User Interface (GUI). IGUANA is completely decoupled from the CMS software, in order to expedite the developments, and very flexible in the sense that the GUIs are highly customizable.

It is possible to work with desktop or web-based GUIs. The remote CMS Control Room at DESY, for instance, uses the web-base mode. This means that IGUANA produces histograms and images on the server and sends them to the web client. This is the suitable mode for remote centers because it allows to use IGUANA GUI without any local installations or software packages.

This web-based system has a drawback for the DQM online event display. This kind of GUI, indeed, is static, i.e. all the components are fast refreshed, but it possible to visualize just the chosen pictures , for instance, of detectors. Consequently, it isn't possible to rotate, to zoom or to rearrange the view of an event.

Due to these reasons, the goal of this project is to provide the CMS Control Room at DESY with an interactive online event display for the DQM Control Room.

This interactive GUI has to be connected with the online stream of data which is produced by the DAQ system. At regular intervals of time, the application reads the stream, randomly catches an event and visualize it. In this way all the data of the picked out event are stored locally, i.e. all the informations about the hits in the sub-detectors can be displayed[2]. Consequently, it is possible to deeply investigate the structure of the event rotating the frame or visualizing particular sub-detectors. After the interval of time, the application skips to an other event deleting the previous one. In this way it is possible to analyze in real time the performances of the detector and to evaluate the kind of events recorded, but the user can't carry on detailed analysis on single events. This is exactly the purpose of the online DQM system, because further analysis are performed later by both off-line DQM and physics analysis.

3.1 ISPY

The first period of my work was spent trying to use the IGUANA event display and facing a lot of troubles because the software was very heavy in terms of both CPU and memory and fragile due to the changing code and data formats. Then I started working with iSpy [3]. It is the new general-purpose event display which doesn't have all the tools of the previous one, but it is more user oriented, i.e. it needs less memory, it is more light and fast in processing events and its code is more stable.

On the contrary of the IGUANA old event display which works with 'root' files [4], iSpy needs files in a particular format, 'ig' files.

3.2 Two Solution

To build up the event display I prepared two different possibilities. The starting point of both of them is P5. This is the computer farm next to the experiment where the stream of events is produced. It is necessary to establish a connection from P5 to DESY through a "double tunnel". Indeed, P5 can't be reached directly from DESY, but only from CERN (fig. 3). So it is needed to pile two 'ssh' connections. This procedure implies that the remote DQM shifter at DESY needs a P5 account.

At this point there are two possibilities:

1. In case the 'ig' files are produced at P5 and the stream is available, then iSpy can run in the online mode with the port redirected to the tunnel;
2. If the 'ig' files are not available, then it is necessary to reach the port where there is the stream of 'root' files. Since this format is not suitable for iSpy,

[2] Online DQM applications use events not yet reconstructed, so there are no tracks or jets.

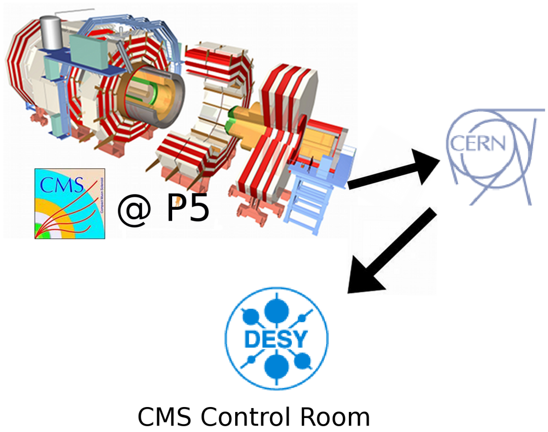

Figure 3: Double ssh tunnel.

it is necessary to convert it in 'ig'. This can be done within the IGUANA framework running a series of instructions contained in a so called "configuration file" that I prepared. This process reads the 'root' files in the stream and provides iSpy with an 'ig' stream. The final result is the same as in the first option.

In the following there are some snapshots of iSpy in the online mode (figg. 4 and 5). It isn't possible to say which of the two will be the default option because these procedures are not completely tested. Last week CMS stopped taking data from cosmic rays and so there was no stream available.

But, at least, I was able to test the second option using a playback server as source of the 'root' stream. A playback server imitates the real online stream from events of previous runs. During this test, the event display was successfully working.

As soon as CMS starts again the runs also the first and simpler option will be tested.

4 CMS Centre Web Site

In parallel with the main project I also worked on the web pages for the CMS centre at DESY [6]. This site is intended to be a tool to organize the activities of the CMS group at DESY, in particular meetings and shifts in the control room. My contribution to the web site is the addition of two separated calendars, one for the meetings and one for the shifts.

The choice on the calendar was very limited because the web server supports only static pages and the most suitable software was Mozilla Sunbird. This is a standalone calendar application and it easily allows to publish calendars, for instance, on ftp server. This calendars can be read and updated with KOrganizer for Unix OS or with iCal for Mac OS, but not with Outlook. The ftp server provides

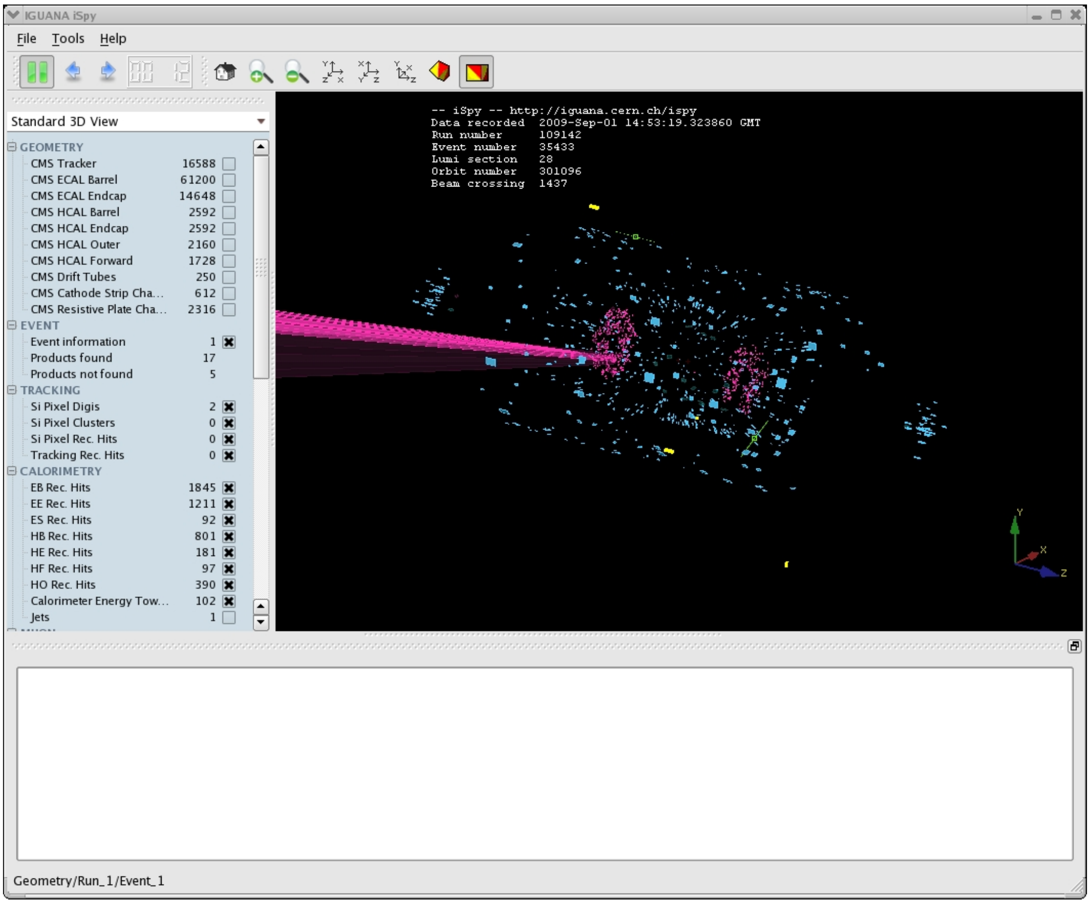

Figure 4: Screenshot of iSpy in the online mode during the test with the playback server. The picture shows the 3D view of the event and it can be rotated and zoomed in/out. On the top is visible the time: after 60 s the event changes. On the left side there is the list of components of the detectors and of hits that can be visualized or hidden.

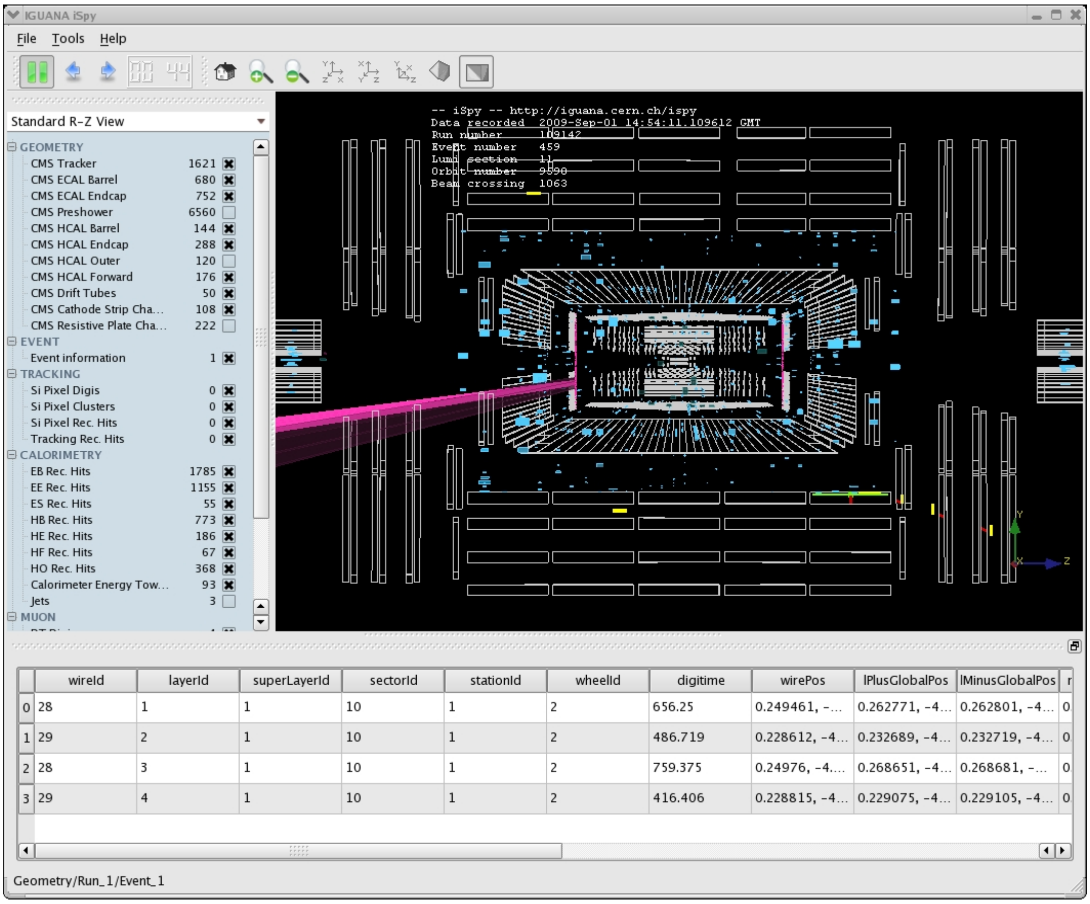

Figure 5: Another picture of iSpy during the test. This is one of the two 2D views (R-Z or R-phi) that are shown for each events. On the bottom is also possible to have quantitative informations about the signals in the sub-detectors.

a easy way to manage the permissions to update the calendars. Indeed, in order to modify the agenda is necessary to substitute the file in the ftp server with the new one and to do that the password for the ftp account is needed.

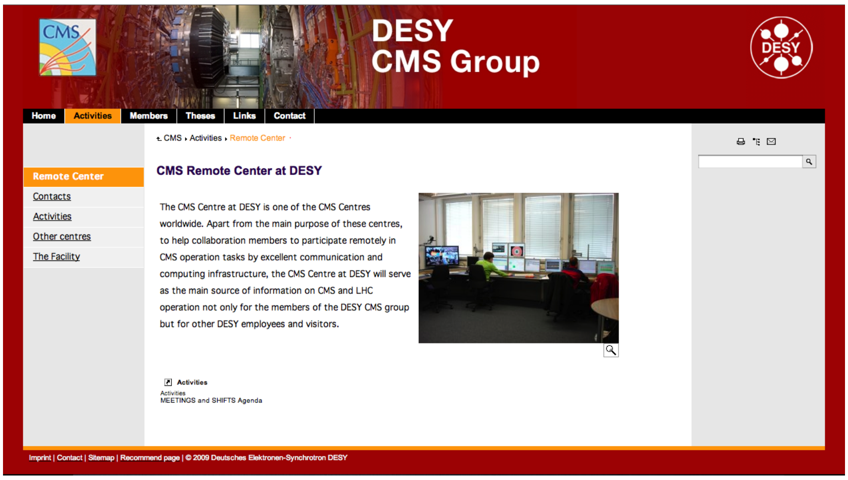

Figure 6: Home page of the CMS remote centre (http://cms.desy.de/e53612/e53619/).

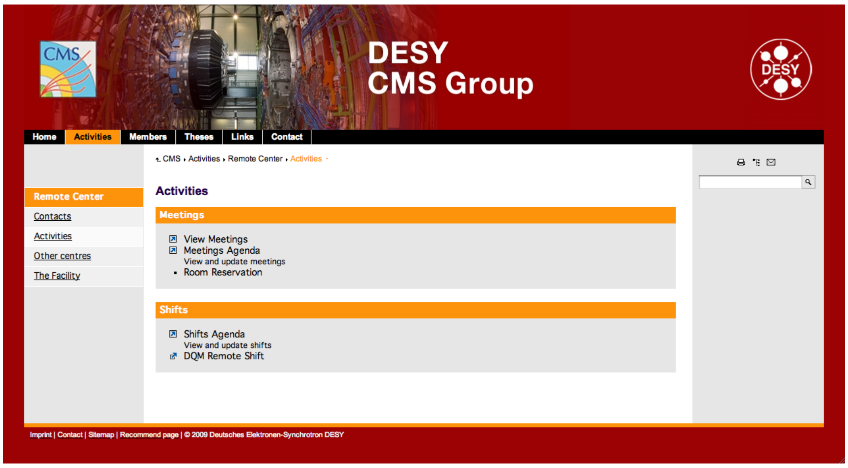

Figure 7: Web page for the activities with meeting and shift calendars (http://cms.desy.de/e53612/e53619/e56480/).

References

[1] CMS Physics Technical Design Report Vol I, CERN/LHCC 2006-001, 2 February 2006;

[2] DQM Twiki pages https://twiki.cern.ch/twiki/bin/view/CMS/DQM;

[3] iSpy documentation http://iguana.web.cern.ch/iguana/ispy/

[4] RooT documentation http://root.cern.ch/drupal/

[5] IGUANA documentation https://twiki.cern.ch/twiki/bin/view/CMS/IGUANA

[6] CMS Centre web page http://cms.desy.de/