

Guillaume Scigala

M1A, 2012/2013

Rapport de stage

LHC@Home portal development

Superviseur : Mr. Peter Jones

Remerciements

Je remercie tout d'abord l'équipe Summer Students du CERN qui m'a permis d'intégrer avec facilité mon département de travail. Mes remerciements vont aussi à mon superviseur Peter Jones qui m'a permis de connaître mon travail, mes collègues et m'a apporté une aide précieuse.

Je remercie également toute ma section au département IT, en particulier David qui fut mon collègue de bureau.

Sommaire

Introduction	3
Rapport	4
Installation	4
Semaine 1	5
Semaine 2	6
Semaine 3	6
Semaine 4	8
Semaine 5	9
Semaine 6	10
Semaine 7 et 8	11
Semaine 9 et 10	11
Conclusion	12
Compte rendu de recherche	13

Introduction

Le CERN (Organisation Européenne pour la Recherche Nucléaire) a été fondé en 1954. Cette organisation, située sur la frontière franco-suisse, est le plus grand laboratoire au monde sur la physique des particules. Le CERN possède 20 états membres du programme de recherche, dont la France.

Le CERN emploie presque 2 400 employés à temps complet, 1 500 à mi-temps et accueille quelques 10 000 visiteurs scientifiques et ingénieurs, représentant 608 universités et centres de recherches de 113 nationalités différentes.

La fonction principale du CERN est de fournir à l'accélérateur de particules ainsi qu'à d'autres infrastructures la recherche en physique en haute-énergie. L'organisation est également le lieu de naissance du World Wide Web. Le site principal situé à Meyrin possède un grand centre informatique avec un puissant centre de calcul, principalement pour des analyses expérimentales de données. Avec comme idée de rendre ces données disponible partout à travers le monde, le CERN est donc historiquement un acteur majeur dans les réseaux informatiques.

Le CERN se décompose en 8 départements :

- HR: Human Ressources.
- FP: Finances & Procurement.
- GS: General Infrastructure Services.
- PH: Physics.
- IT: Information Technology.
- BE: Beams.
- TE: Technology.
- EN: Engineering.

Durant la période estivale, du 24 Juin au 30 Août 2013, je fus affecté au département IT par l'intermédiaire du programme «Summer Student». Ce département est divisé en plusieurs groupes et sections : CF (Computer Facilities), CIS (Collaboration and Information Services), CS (Communication Systems), DB (Databases Services), DI (Department Infrastructure), DSS (Data & Storage Services), OIS (Operating System & Infrastructure Services), PES (Platform & Engineering Services) et SDC (Support for Distributed Computing). Je fus affecté au groupe PES. Parmi ce groupe, différentes équipes travaillent ensemble: mon équipe suivait la dénomination IS (Grid Information Service).

Ma mission fut la création d'un nouveau portail pour le programme LHC@home. Le CERN réalise un projet informatique avec des volontaires, le LHC@home, utilisant le framework BOINC permettant de séparer les simulations aux volontaires du monde tout entier. Le projet fut donc de déployer un portail web aux utilisateurs volontaires de LHC@home à l'aide de Drupal, gestionnaire de contenu web. Une partie du travail fut donc d'établir les liens entre BOINC et Drupal pour une meilleure intégration entre eux.

Rapport

Installation

Lors de mon arrivée au CERN, une présentation a été donnée par les membres organisateurs du programme «Summer Student». L'explication du fonctionnement du site m'a été faite, ainsi que son but ainsi que son organisation interne. M'ont également été fourni l'adresse du secrétariat auquel j'étais affecté pour ma mission. Sur la période du 3 Juillet 2013 au 9 Août 2013, chaque matinée, 4 heures de conférences sur les travaux du CERN sont ainsi faites, me donnant le choix soit de réaliser ma mission dans le département IT ou de suivre celles-ci. Cette première introduction a donc eu lieu le lundi 23 Juin 2013 à 9h. Les participants, environs 50 personnes, étaient tous étudiants et venaient du monde entier, principalement d'universités de Physique.

Postérieurement à cette conférence, mes codes d'accès internet ainsi que ma carte d'accès me furent délivrés au poste de garde du site pour me permettre de me mouvoir en toute commodité durant les 10 semaines de ma mission. Après cette étape, le secrétariat m'indiqua qui était mon superviseur. Ce dernier, Pete, me fit une visite du bâtiment avec toutes les personnes avec qui je serais potentiellement en contact. Il m'indiqua également où se trouvait mon bureau.

Mon bureau, au bâtiment 31, était donc partagé avec deux personnes : Jasiak et David. Ces deux personnes se montrèrent très serviables tant sur l'installation de mon poste de travail que sur les subtilités du CERN sur son mode de fonctionnement, notamment pour la partie réseau.

Mon poste de travail fut donc un ordinateur avec un OS de type Linux, SLC6. Cet acronyme signifie Scientific Linux CERN, cette distribution est fournie par Red Hat, et est ainsi spécifique aux besoins du CERN. La connexion à ce poste de travail se fait par l'intermédiaire de mon compte CERN. En effet ce compte détient tous les accès sur le site, où j'ai dû enregistrer ce poste de travail, mais également mon ordinateur portable et mon téléphone grâce aux adresses MAC pour pouvoir accéder au réseau.

D'un point de vue humain, tout échange au sein de l'équipe se fait en langue anglaise, et tout rapport écrit également. Concernant le suivi de ma mission, le département IT a mis en place TWiki, une plateforme de collaboration directement incorporée au CERN. Voici un récapitulatif des personnes pouvant m'aider pour ma mission:

Nom	Rôle	Coordonnées	Courriel
Hoimyr Nils	Chef de section	IT PES-IS 31 2-001	Nils.Hoimyr@cernSPAMNOT.ch
Jones Pete	Superviseur	IT PES-IS 31 2-005	Peter.L.Jones@cernSPAMNOT.ch
Marquina Miguel	Liens entre projets	IT DI 600 R-011	Miguel.Marquina@cernSPAMNOT.ch
Gonzalez Alvarez Alvaro	Copilote	IT PES-IS 31 1-018	Alvaro.gonzalez.alvarez@cernSPAMNOT.ch
Segal Ben	LHC @ Home 2.0	IT 600 R-014	B.Segal@cernSPAMNOT.ch

Figure 1: tableau coopératif

Semaine 1

Le début de semaine fut donc marqué par l'installation sur les deux premiers jours de ma mission. Après avoir ainsi pu m'installer convenablement, avec un matériel fonctionnel et prêt à l'emploi, je suivis le planning qui m'avait été donné sur TWiki:

- Suivre la procédure d'installation CERN.
- S'installer dans le bureau.
- Obtenir un compte CERN, se connecter et se familiariser avec l'environnement informatique CERN.
- Devenir membre des nécessaires BOINC et IT egroups.
- Prendre connaissance de TWiki.
- Être familier avec l'infrastructure BOINC et les projets BOINC au CERN.
- Être familier avec les différents sites web BOINC qu'il y a au CERN.

Les quarts premiers points ayant été réalisés, je me suis attelé à prendre connaissance de Twiki. Cet outil permet de collaborer avec toute l'équipe par l'intermédiaire du domaine cern.ch. En effet chaque note, compte-rendu et différents ou autre document ayant été préalablement entré dans le système est automatiquement lisible par tous, permettant une collaboration plus fructueuse.

Concernant l'infrastructure BOINC, j'ai ainsi pris connaissance de ce logiciel. Ce dernier (Berkeley Open Infrastructure for Network Computing) est un système middleware pour volontaires et pour le GRID Computing. BOINC à initialement été développé pour le programme SETI@home avant de devenir une plateforme utile pour d'autres applications comme les mathématiques, la médecine, la biologie moléculaire, le climat, l'espace ou ici l'accélérateur de particule. Le but de BOINC est de permettre aux chercheurs de disposer de la puissance processeur de différents PC à travers le monde.

Au sein du CERN, BOINC est utilisé pour deux projets : «Test4Theory»(T4T) et «SixTrack».

- ▶ «Test4Theory» est un projet permettant la simulation de collision. En effet, par l'intermédiaire de l'infrastructure BOINC, ce dernier exploite la technologie de machine virtuelle permettant aux utilisateurs de contribuer à la recherche de nouvelles particules fondamentales au LHC du CERN.
- ▶ «SixTrack» quant à lui est un autre projet du CERN, permettant d'étudier le fonctionnement dynamique des collisions dans le LHC. Cela traduit ainsi les améliorations possibles de l'accélérateur pour pouvoir comprendre les lois fondamentales de l'univers.

Pour le dernier point m'ayant été attribué, différents sites internet sont en effet disponibles depuis la page <http://lhathome.web.cern.ch/LHCathome/>. En étudiant celle-ci, j'ai pu apprendre quels étaient les différents objectifs du CERN concernant le projet LHC@home

Ayant été efficace sur ces points, à partir du mercredi je fus invité à créer un site web de test pour me permettre de me familiariser avec l'outil Drupal. Ce dernier, étant un gestionnaire de contenu web, est directement disponible à l'usage après une requête au service internet du CERN. Il m'a donc été permis de tester Drupal, modelant à ma guise le site de test pour me rendre plus efficace lors du modelage du vrai site internet dans les semaines à venir.

Semaine 2

Au début de cette semaine me furent confiées deux tâches concernant le site en cours de développement LHC@Home:

- Vérifier tous les liens du site, et les modifier en conséquence.
- Rassembler les deux sites internet des deux projets distincts sur ce site de développement pour ainsi permettre à l'utilisateur d'avoir une vue d'ensemble des différents projets.

Après avoir réalisé ces deux tâches, mon superviseur m'indiqua qu'il était préférable que je m'attèle à une autre tâche: la séparation sur le même site des contenus des deux projets. Pour ce faire, Pete me confia un cours sur Drupal, pour ainsi me permettre de maîtriser toutes les subtilités du CMS Drupal (Content Management System). L'idée était d'utiliser une spécificité de Drupal, la Taxonomie, pour pouvoir en différencier les contenus. Pour le restant de la semaine, je fus donc à l'entraînement grâce à ce cours sur mon site de développement, essayant également de modifier les différents thèmes possibles pour le futur site internet.

Parallèlement à cela, les conférences «Summer Student» au CERN commencèrent. Celles-ci ont pour but d'expliquer aux étudiants du programme quels sont les travaux réalisés au CERN. Voici une liste des différentes conférences, rassemblées en trois points, données lors de ce programme:

- Physique expérimentale: Accelerators, Detectors, Hadron collider physics, Flavour physics, Neutrino physics, Astrophysics, Antimatter, Nuclear physics, Heavy ions, etc.
- Physique théorique: Quantum field theory, Standard Model, Beyond the Standard Model, Cosmology, Astroparticle physics, etc.)
- Autre: Statistics and Data mining, Electronics, DAQ systems, Medical physics, ROOT, etc.

Toutes ces conférences ont eu lieu les matins du 3 Juillet 2013 au 9 Août 2013.

Semaine 3

Cette semaine je commença à implémenter différentes méthodes pour pouvoir séparer, au sein de mon propre site de développement Drupal, les différents projets du CERN. Ainsi la Taxonomie me fut d'une grande utilité, et ainsi les trois projets furent rassemblés. Voici ci-dessous une copie de mon rapport de fin de semaine concernant ce site:

*You can see a developpement version with Bootstrap apparence for the [lhc@home](https://dev-test-31-1-003.web.cern.ch/) project here : <https://dev-test-31-1-003.web.cern.ch/>
This website merge the three projects of CERN in a unique website : [SixTrack](#), [Test4Theory](#) and [Beauty](#). The selection of the main project is done on the front page in the main menu.
With Taxonomy, the content is splitted into the three project. For example, if you are in the [Test4Theory](#) project, the URL look's like <https://dev-test-31-1-003.web.cern.ch/test4theory/>.
The Taxonomy allow to know where we are, and so we can modify the blocks in consequence. They are two modification whose depends of the projet :*

- *In the primary bar (in the left), we have a link to a redirection to the other project.*
- *In the menu bar (in the top), they are modification of the menu available.*

*In the secondary bar (in the right) are displayed all the recent content of the three projects, with an independence of where you are in the website. They are a RSS feed too with the news of BOINC.
The CERN bar isn't displayed because it's not optimize with the Bootstrap apparence.
Quick links are available in the left bar, is avoid a doubloon of the menu like the [lhc@home](#) dev site. This block is displayed independammently of the position in the website.
The login is on the left bar too, and the account access in the menu bar.*

Guillaume Scigala, 11 July 2013

Durant cette même semaine, une session tutorial me fut donné concernant le logiciel ROOT. Ce dernier fourni un framework orienté objet avec toutes les fonctionnalisés requises pour traiter une quantité importante de données et ce avec efficacité. Ce logiciel est donc utilisé par le CERN pour permettre aux physiciens d'utiliser les données provenant des collisions du LHC, et d'ainsi faire avancer leurs différents travaux concernant les particules élémentaires. Voici ci-dessous un exemple de graphique obtenu grâce au logiciel ROOT:

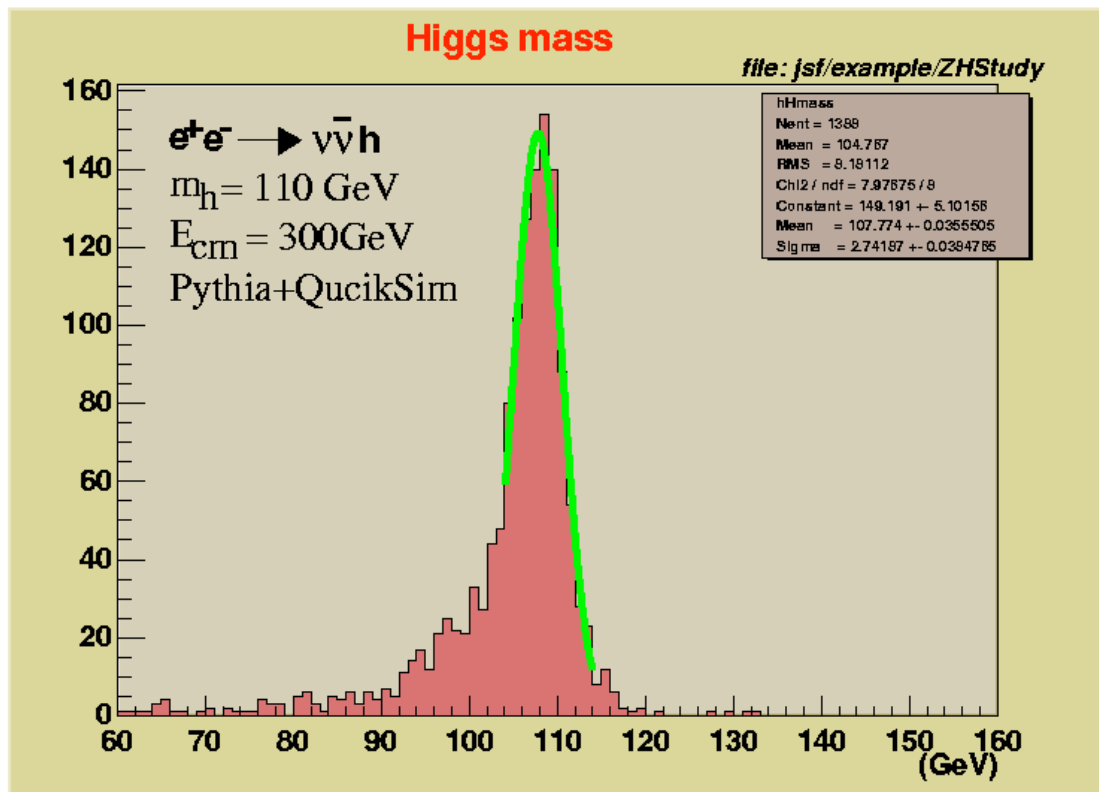


Figure 2: Masse du boson de Higgs grâce au logiciel ROOT.

Cette semaine furent également données les premières conférences «Openlab». Ces dernières traitent de l'informatique et contrastent avec les conférences matinales. Voici ci-dessous la liste des différentes conférences:

Lecture 1: General Security (S. Lueders)	Monday, 8 July, 2013 - 16:00 to 17:30
Lecture 2: Software Security (S. Lopienski)	Thursday, 11 July, 2013 - 16:00 to 17:30
Lecture 3: Web Application Security (S. Lopienski)	Tuesday, 16 July, 2013 - 16:00 to 17:30
Lecture 4: Cloud Computing in Large Computer Centers (B. Rodrigues Moreira)	Thursday, 18 July, 2013 - 16:00 to 17:30
Lecture 5: Physics Computing	Tuesday, 23 July, 2013 - 16:00 to 17:30
Lecture 6: Machine Learning for LHCb with Yandex MatrixNet (A. Ustyuzhanin)	Thursday, 25 July, 2013 - 16:00 to 17:30
Lecture 7: Worldwide LHC Computing Grid Overview (F. Furano)	Tuesday, 30 July, 2013 - 16:00 to 17:30
Lecture 8: TBD	Tuesday, 6 August, 2013 - 16:00 to 17:30
Lecture 9: Data Reliability at CERN and Ideas on How to Improve it (A. Pace)	Tuesday, 13 August, 2013 - 16:00 to 17:30

Lecture 10: Oracle Databases at CERN (E.Grancher)	Thursday, 15 August, 2013 - 16:00 to 17:30
Lecture 11: TBD	Tuesday, 20 August, 2013 - 16:00 to 17:30
Lecture 12: Systemtap : Patching the linux kernel on the fly (T.Oulevey)	Thursday, 22 August, 2013 - 16:00 to 17:30
Lecture 13: Control System Cyber Security (S. Leuders)	Tuesday, 3 September, 2013 - 16:00 to 17:30

Figure 3: liste des différentes conférences «Openlab»

Ainsi deux conférences furent données cette semaine, l'une sur la sécurité du CERN et l'autre sur la sécurité logiciel.

- Concernant la première, le CERN dispose d'une politique sécuritaire spécifique. En effet l'organisation est à la fois baignée dans le milieu universitaire avec des membres du personnel arrivant du monde entier, mais elle est également sous le feu d'attaque informatique mondiale. Ainsi le CERN permet aux utilisateurs de disposer d'un compte, d'autoriser ses périphériques, mais confie la responsabilité sécuritaire à chaque membre du personnel. Le service de sécurité dispose ainsi d'une vue d'ensemble du réseau, apportant différents services à son personnel comme des bases de données, des logiciels contrôlés et sécurisés pour une utilisation professionnelle.
- La seconde conférence portait sur la notion de sécurité logiciel. En effet, chaque code informatique peut posséder différentes failles comme une exécution d'un code malicieux lors d'une saisie de commande par l'utilisateur: cette notion de sécurité est ainsi primordiale. Elle est également à mettre au coeur de la conception d'un logiciel et intrinsèque à chaque étape de développement, allant de la définition des spécifications aux tests.

Semaine 4

Durant cette semaine, ma mission fut d'étudier une intégration entre le projet BOINC et le site internet sous Drupal dans le but de créer un forum. En effet, BOINC dispose d'une base de données dans laquelle chaque utilisateur possède son score sur le projet ainsi que différentes autres caractéristiques intrinsèques au projet BOINC. Drupal disposant également de son propre mode de fonctionnement, notamment pour le classement des fichiers sur le serveur, je dus donc étudier quelle serait la possibilité de regrouper ces deux bases de données.

Je fus ainsi confronté à quatre problèmes:

1. La source de BOINC permettant l'intégration était manquante, mais à l'intérieur de celle du client non compilé était présent un fichier Drupal.
2. Un autre projet ayant déjà réalisé cette intégration, Einstein@home, je découvris par mes recherches qu'un module spécifique avait été développé pour Drupal permettant l'intégration. Je demandai donc d'avoir une copie de celui-ci.
3. Pour pouvoir intégrer BOINC, je devais également exécuter certaines commandes dans le répertoire du site internet, sur le serveur du CERN. Cette opération était impossible avec les droits utilisateurs qui m'étaient alloués.
4. Enfin j'avais besoin d'une copie de la base de donnée du projet BOINC. Ceci pose un problème: sachant que le site internet regroupe différents projets, quel doit être la base de données importées? Peut-être la possibilité de différents forums doit être envisagée.

Durant cette semaine, je continuais de suivre les conférences matinales, mais également celles destinées au département informatique les mardis et jeudis après-midi.

Semaine 5

Cette semaine fut la continuité de la précédente, avec en particulier la préparation sur le site de développement la mise en place du thème «Bootstrap». En effet ce dernier a été retenu pour être le thème par défaut du futur site internet. La préparation c'est donc déroulée par la préparation des contenus avec la Taxonomie, permettant de séparer les différents projets sur le site internet. Ceci consista à ajouter ce champs à tous les contenus déjà présents.

La présentation fut également faite du mode de fonctionnement informatique au CERN. Ce système repose sur deux entités : LxPlus et LxBatch. La première est l'interface de communication des personnes du CERN, avec un espace disque alloué. La seconde permet le traitement des données provenant de l'accélérateur de particules. En effet le CERN a séparé en deux son traitement informatique des données: l'un utilisant la Grille, et l'autre LxBatch.

La Grille est un réseau communautaire de traitement de données, avec de nombreuses universités et pôles de recherches partenaires. Celle-ci permet d'exporter les calculs à l'extérieur du CERN.



Figure 4: Aperçu de la Grille

Pour pouvoir découvrir de nouvelles particules, le seul fait de réaliser des collisions de proton au sein du LHC ne suffit pas: il faut savoir se que l'on cherche. Ainsi, d'après le Modèle Standard, de nombreuses simulations sont effectuées, comprenant trois fois plus de calculs que ceux de l'exploitation des données provenant du LHC. La Grille possède le rôle de simulateur, toutes les données provenant du Modèle Standard sont ainsi exploitées pour simuler la future découverte. L'exploitation des données se fait quant à elle sur place, grâce à LxBatch.

Ces deux solutions sont ainsi combinées pour pouvoir découvrir et comprendre le monde de la physique des particules.

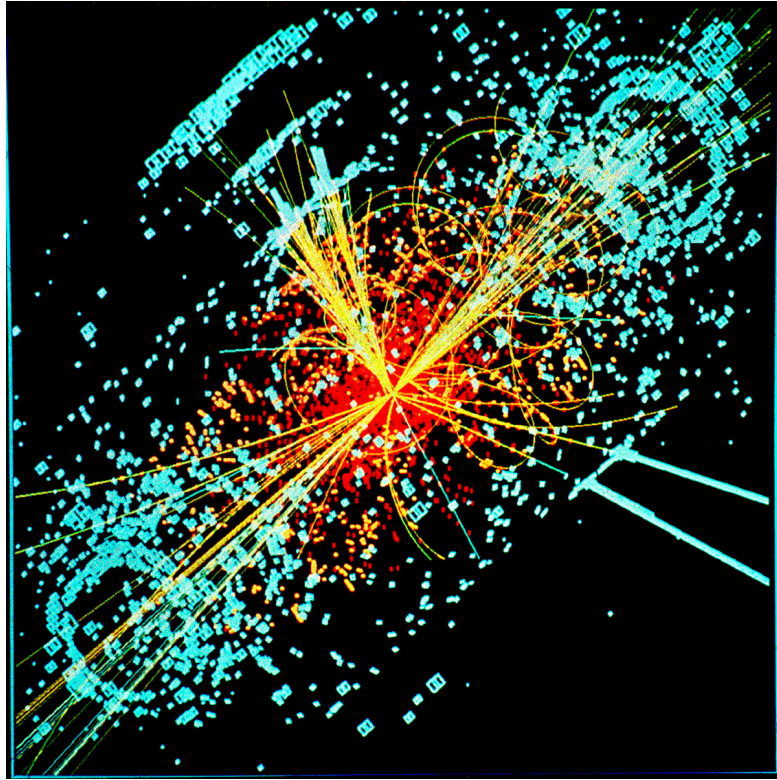


Figure 5: Simulation du boson de Higgs

Semaine 6

Cette semaine a été marquée tout d'abord par la modification des apparences suivant le projet voulu sur le site internet sous Drupal. En effet, après une réunion, il a été décidé que la page d'accueil du site serait un portail regroupant les différents projets, permettant ainsi de réunir les différents sites en un seul, ce dernier regroupant les projets sous différentes apparences. Je commençai donc la modification du site de développement du LHC@Home pour correspondre aux attentes.

Durant cette semaine je réalisai également une expérience, la chambre à nuage. Celle-ci permet de visualiser les particules grâce à un mélange chimique créé sous cloche. On peut ainsi distinguer par des traits dans le nuage les différentes particules arrivant de l'espace. Cette expérience est donc sensible aux particules émanant suivant l'axe horizontal.

Je fis également la visite de l'expérience CMS (Compact Muon Solenoid), un détecteur du LHC.

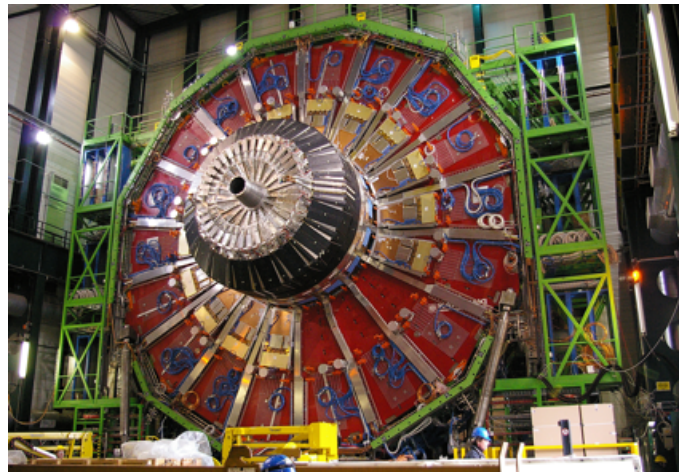


Figure 6: Détecteur CMS

Semaine 7 et 8

Durant ces deux semaines, ma tâche fut la mise en production du site internet.

Au départ, il a été défini quels étaient les différents contenus voulus pour chaque projet. Ainsi ceci a permis de séparer le site internet plus simplement, avec aucun conflit notamment sur les différents menus affichés. La création de la page de garde fut également travaillée, avec un diaporama sur cette page. Celle-ci allait donc être le portail du site, dirigeant vers les différents projets actuels et futurs de LHC@home.

Sur ce site a également été implanté un menu pour les écrans mobiles, permettant l'accès au site depuis n'importe quel appareil. Celui-ci procède à un changement de couleur en fonction du projet en cour de consultation. Enfin, chaque lien a été vérifié pour s'assurer du bon fonctionnement du site internet.

Parallèlement à cela, j'ai eu la possibilité de visiter les infrastructures de LINAC (Accélérateur Linéaire de proton), du centre de calcul ainsi que de l'expérience ATLAS.



Figure 7: Linac 2

Pour d'avantage de renseignements sur CMS, ATLAS ou LINAC, consulter cette page : <http://home.web.cern.ch/fr/about/>

Semaine 9 et 10

Le travail effectué ces deux dernières semaines fut d'abord la démonstration à l'équipe de LHC@home le nouveau site internet regroupant leurs activités. Cela amena différentes questions, notamment la place des éditeurs du site internet. J'ai ainsi travaillé à permettre à ces derniers d'accéder facilement aux commandes d'édition et d'ajout de contenus.

Je réalisai également une démonstration sur le site de développement de la possibilité de sélection de différentes langues pour l'utilisateur du site, Drupal étant un CMS pouvant traiter différentes langues. Je fis donc une description de la procédure à suivre pour mettre en place cette fonctionnalité, mais également de la manière d'ajouter différents contenus traduits ou non.

Conclusion

Au terme du contrat de travail, le projet défini à la signature de ce dernier fut mené à bien, avec une construction logique et claire du site internet. Concernant l'avenir, les futurs projets se rattachant au LHC@home posséderont dorénavant une meilleure visibilité, avec un hébergement sur un site et donc serveur commun. Ma mission a donc permis cette approche de cohésion, pour une communication vers l'extérieure du CERN plus aisée.

Cette expérience d'étudiant d'été me fut enrichissante d'abord d'un point de vue humain grâce à l'intégration dans une équipe fonctionnelle et organisée du CERN, montrant comment et pourquoi fonctionne cette organisation internationale. Techniquement, mon travail au CERN m'a permis d'acquérir une expérience sur la création, le fonctionnement et la maintenance d'un site internet, et plus particulièrement sur les CMS. Les différentes conférences données durant l'été m'ont donné la possibilité de connaître quels étaient les différentes problématiques du CERN à savoir la recherche de particules élémentaires, la compréhension de l'univers ou encore le mode de fonctionnement des accélérateurs de particules. Ces conférences m'ont également permis de prendre conscience du traitement des données acquises par les détecteurs, le challenge de la fiabilité et de la stabilité de celles-ci.

D'un point de vue local, le poste que j'ai tenu durant ces 10 semaines fut le développement de ce nouveau portail, mes responsabilités furent ainsi transférées à chaque dirigeant de chaque projet du LHC@home à savoir Test4Theory, SixTrack et Beauty. L'équipe à laquelle j'ai appartenu a ainsi joué son rôle dans le succès de ma mission de part sa présence, mais également son expertise et son savoir-faire sur la question de site internet. Le CERN étant une organisation de recherche, l'état d'esprit global est ainsi intéressant avec la notion de compétence: en effet chacun est spécialiste dans son domaine et chaque problème peut être résolu en allant prendre certains renseignements clés à un collègue de section ou de groupe.

Le CERN m'est donc apparu non seulement comme une organisation avec une grande mixité culturelle de part ces différentes nationalités, mais également comme un lieu où chacun se voit confié une tâche qui n'est non plus unique et individuelle mais globale et enrichissante pour la personne, et in fine pour la communauté scientifique toute entière.

Compte rendu de recherche

Pour pouvoir trouver ce stage, j'ai consulté le site web du CERN <http://home.web.cern.ch>. Ainsi j'ai pu découvrir différents programmes pour étudiants que l'organisation proposait, dont le «Summer Student Programme». J'ai donc postulé, ayant fourni CV, lettre de motivation ainsi que deux lettres de recommandation de la part de deux de mes professeurs, Mr. Gemal et Mme. Chenoune.