

**PROPOSITION DU SUJET DE MEMOIRE
AUDITEUR DU CENTRE PARIS**

Nom et prénom de l'élève ingénieur: Eléonore HOLLE

PROFESSEUR RESPONSABLE NATIONAL DE LA SPECIALITE: Yann POLLET
INTITULE DU SUJET : Etude et mise en œuvre d'une interface Observatoire Virtuel permettant l'accès à la base de données HFC (Heliospheric Feature Catalogue).

RESUME (description du contexte, de votre fonction et du projet) :

L'Observatoire virtuel (OV) européen HELIO (<http://www.helio-vo.eu/>) fournit des services en ligne dans le domaine de la physique solaire et interplanétaire. Une partie de ces services est actuellement hébergée et maintenue par l'Observatoire de Paris, en collaboration avec le centre régional d'expertise "Paris Astronomical Data Centre" (PADC).

L'équipe HELIO de l'Observatoire souhaite rendre l'un de ces services, le Heliophysics Feature Catalogue (HFC), interopérable avec des portails et outils OV au travers du protocole EPN-TAP, développé dans le cadre du projet Europlanet.

Dans une première partie, il s'agira de prendre connaissance de l'architecture actuelle et de se familiariser avec le protocole EPN-TAP afin d'avoir une vision claire de l'architecture cible.

Plusieurs solutions sont envisageables dont les deux grandes lignes sont:

Garder le gestionnaire MySQL actuel pour la base du HFC, et utiliser une bibliothèque Java du Centre de Données astronomiques de Strasbourg (CDS), pour développer l'interface compatible avec le protocole EPN-TAP. Si l'on souhaite par ailleurs autoriser des requêtes de type spatial, il faudra vérifier que le module `mysql_sphere0.1` offre les mêmes fonctionnalités que `pgSphere` de PostgreSQL.

Changer le gestionnaire de la base pour PostgreSQL et utiliser un framework python, nommé `DaCHS`, pour réaliser l'interface compatible avec le protocole EPN-TAP. Cela nécessitera une modification du code de toutes les interfaces déjà existantes qui interrogent pour l'instant spécifiquement une base MySQL. Il manque une couche d'abstraction qui nous aurait permis de changer de gestionnaire de base de données sans avoir à toucher ces codes. Par ailleurs, le protocole EPN-TAP utilise un nommage standard des données exposées. Il sera peut-être utile dans ce cas d'adapter les noms des données de la base et cette couche d'abstraction servirait également d'adaptateur pour la consultation de ces données via le protocole EPN-TAP.

Il s'agira d'étudier la faisabilité, les bénéfices/risques des différentes architectures possibles.

A partir de cette première analyse, il s'agira de proposer des solutions techniques et un plan de développement adaptés, tenant compte des répercussions sur l'architecture actuelle du service.

La troisième phase du projet devra voir la mise en œuvre pratique, comprenant le développement, le test et la validation de la nouvelle infrastructure avec éventuellement la supervision d'un stagiaire qui réaliserait une partie du développement.

PLAN DE TRAVAIL :

15 janvier - fin février 2016 : Appropriation de l'existant et de la cible voulue:

De nombreux documents sont à lire et à comprendre (tous sont en anglais):

Le **protocole TAP** (Table Access Protocol) dont dérive le protocole EPN-TAP

Le **protocole EPN-TAP** qui constitue notre cible.

Le **Framework python GAVO DaCHs** développé par l'équipe GAVO en Allemagne.

La Base de données **HFC**.

Le **TAPRegExt** : le standard d'encodage XML des métadonnées pour être compatible avec le protocole TAP.

Un document qui décrit la structure minimale à respecter pour une **interface** qui permet des échanges REST ou SOAP dans l'environnement IVOA (International Virtual Observatory Alliance)

Le **Registry** (IVOA Registry Interface) qui décrit les moyens de recherche et de récupération des données et services pour les applications de l'Observatoire Virtuel.

Une Comparaison MySQL- PostgreSQL axée sur nos besoins cibles.

Les deux "Foreign Data Wraper" implémentés pour permettre l'utilisation du framework DaCHs, compatible EPN-TAP.

Le langage de requête spécifique qu'il faut utiliser pour les données astronomiques : **ADQL** (Astronomical Data Query Language)

Le pattern "**Universal Worker Service Pattern**" utilisé pour implémenter un service asynchrone.

Les résultats de requête sont des **VOTable**. (Etudier donc ce qu'est une VOTable).

Le **protocole SAMP** (Simple Application Messaging Protocol) adapté lui aussi aux données astronomiques.

Fin février - 15 mars: Etude de la faisabilité et des risques de chaque solution.

Décision de l'axe majeur à prendre.

Choix techniques et répercussions sur l'architecture existante.

Définition précise du plan de développement.

15 mars - 15 juin : plusieurs itérations comprenant chacune : développements des composants, tests unitaires, tests d'interface entre composants ou stubs, rédaction de la documentation spécifique à chaque composant

15 juin - 15 juillet : Tests d'intégration de l'ensemble des composants, déploiement, conclusions, documentation finale.

15 juillet - fin juillet : perspectives :

Utilisation de cette expertise acquise pour adapter d'autres services HELIO de l'Observatoire de Paris au protocole EPN-TAP.

Août - septembre (moins 15 jours vacances) : rédaction du mémoire.

Octobre 2016: Préparation de la soutenance.

AVIS DU PROFESSEUR:

Date et signature :