

Projet de Compte-rendu de la réunion de suivi du CDS du 15 décembre 2009

Présents: M. Allen, L. Arbousse, T. Boch, F. Bonnarel, C. Bot, L. Cambresy, S. Derriere, P. Fernique, F. Genova, S. Lesteven, C. Loup, F. Ochsenbein, A. Preite Martinez, A. Schaaff, A. Siebert, B. Vollmer, M. Wenger, H. Wozniak.

Informations générales

- Action : mettre en place des réunions régulières pour les chercheurs à partir du mois de janvier 2010, afin notamment de faire le tour des tâches de services et des questions de stratégie scientifique.
- 2 préoccupations fondamentales: la fiabilité des services/qualité des services d'une part, et une préoccupation de type "assurance/produit" soulevée par R. Bellenger et P. Puget, d'autre part.
- Equilibre VO / autres activités:
En ce qui concerne les projets européens, un projet d'une année a été rapidement préparé avec les partenaires italiens, anglais ainsi qu'avec l'ESO en vue de maintenir les "Technology Forum", un « Hands-on » et le Science Advisory Committee. Ce projet comprend seulement des missions. Il devrait y avoir un autre appel d'offre sur les "Data Repositories" l'année prochaine, auquel le CDS répondra.
- Bugs: retours suite aux différentes réunions (Madrid,...): concernant Aladin, des bugs ont été évoqués et des propositions d'évolution faites. Il a aussi été noté qu'Aladin a corrigé beaucoup de bugs signalés aux mois de mars et juin.
- Le Conseil Scientifique du CDS est en cours de recomposition et une réunion se tiendra avant l'été 2010. Ce sera toujours un Conseil international composé de 6 membres français, 6 membres non français, il inclura 1 représentant du CNES, 1 représentant du Ministère, 1 invité de l'Université ainsi qu'1 invité de l'INSU. C'est un Conseil où sont présentées les activités du CDS, les services, la stratégie.
- Action: le bilan de l'utilisation des services en 2009 (nombre de hits, nombre de downloads pour Aladin). L'école du mois de janvier, EuroVO-AIDA Hands-on Workshop qui se tiendra du 25 au 28 janvier 2010, sera utile pour étudier l'utilisation et l'ergonomie: il y aura des groupes spécifiques et il faudra aller voir de plus près dans ces différents groupes comment cela se passe. Parmi les statistiques intéressantes: sur les 4 derniers jours, le flux RSS a été demandé par plus de 600 IP.
- Pour assurer la fiabilité du point de vue du matériel, il y a un problème lié à l'absence de redondance au sein de la salle informatique. Il faut plusieurs compteurs électriques; on pourrait aussi avoir 2 salles informatiques pour avoir tous les services en double (il faut penser aux pannes électriques et aux pannes de clim). Il y a également un gros souci de liaison entre les bâtiments Est et Sud (où la puissance n'est pas suffisante). Il faudrait donc avoir 2 sources électriques différentes et 2 salles informatiques indépendantes ; le dossier est poussé par H. Wozniak. Discussion en cours avec la Direction du Patrimoine immobilier pour le maintien de deux compteurs distincts, un chez nous et un au Planétarium.
- Gestion du compte question
La gestion est plus classique sous thunderbird ; c'est bien mais ce n'est pas parfait.

- Permanence pendant les fêtes de Noël ? Action menée par M. Wenger en liaison avec J.-Y. Hangouet.
Des redémarrages et des transferts de machine sont prévus jusqu'au mardi 22 décembre. Il n'y aura qu'une clim à fond (et non les deux à moitié, ce qui risquerait de faire tout sauter) jusqu'à ce que les travaux soient finis, avec un switch en cas d'arrêt.
- Matériel: l'installation d'une des machines, la nouvelle CDS Portal pour permettre le démarrage du service dès que possible, est à mener à bien de manière urgente, en raison de la revue AIDA qui se tiendra en mars.
- Direction du CDS
H. Wozniak a demandé à F. Genova de continuer à diriger le CDS (son mandat arrivant à expiration cet été 2010) pendant la période d'incertitude qui entoure la fusion des 2 OSU.
Puis il y aura un appel d'offre pour la direction du CDS.

Simbad

Bibliographie

- Personnel
Réorganisation importante liée aux départs à la retraite de S. Borde et P. Dubois ainsi qu'au poste obtenu par M. Neuville.
Le travail qu'effectuait P. Dubois est maintenant réparti entre L. Cambresy, A. Siebert, C. Bot, C. Loup et B. Vollmer.
- Réunions tous les 15 jours organisées par M. Brouty pour les « djinistes », afin de parler notamment des problèmes d'identification. Il faut assurer le contact avec G. Chassagnard.
- Bilan journaux
Le volume de données à traiter augmente de façon significative:
 - MNRAS a doublé en 8 ans
 - A&A a augmenté de 20%
 - ApJ a aussi augmenté
 En 2008, 11.000 références par an sont rentrées et ce sont 42 journaux qui sont traités. En ce qui concerne le suivi des grands journaux, ça va encore. D'autres journaux plus petits restent à la traîne...
- Toute la documentation nécessaire aux documentalistes a été refaite et complétée. A. Oberto s'occupe de tout ce qui est Mise à jour ; A. Eisele a repris la documentation dont s'occupait S. Laloë. M. Neuville et E. Perret ont apporté leur contribution. C. Bonnin, qui part fin février, travaille sur la documentation de DJIN. Tri « de pertinence » des références attachées aux objets par rapport à la position des noms d'objets dans le texte. Ceci est fait sur tous les titres en remontant au début de Simbad (en cours de vérification par S. Borde).
- A&A
Il n'y aura plus de pagination dès le mois de janvier, mais des numéros d'articles. Le bibcode va donc évoluer.
- F. Ochsenbein, S. Lesteven, M. Brouty et F. Woelfel participeront à LISA VI et vont présenter 2 papiers.

Contenu

- Les tables
Les vieilles listes en retard ont été terminées (tri + entrée). L'équipe se concentre sur les années 2007 à 2009 mais est tributaire de VizieR. E. Perret passe 2 jours par semaine pour aider à rattraper le retard pris sur VizieR.
- Tri des listes à entrer dans SIMBAD
Les chiffres des entrées/sorties VizieR restent stables. Il y a un peu de retard dans le tri: 150 listes restent à trier.
Le traitement des listes en priorité 1 est constant, ce qui est bien pour les documentalistes. Les priorités 2 s'accumulent. La mise en exploitation du Portail, qui permettra de chercher dans tous les services en même temps, est une amélioration fondamentale qui va assurer un accès beaucoup plus facile au contenu de ces listes.
- Combien de listes sont susceptibles de passer dans Simbad, par année de parution ?
457 pour 2007 (c'est le total de ce qui est trié en =g=): 37% en priorité 1 ; 35% en priorité 2.
12 à 14% des listes sont traitées de manière informatique.
Il manque environ 30% de listes.
- Le nom d'objet principal dans la première fenêtre d'affichage de Simbad est souvent abscons ; un gros travail est en cours afin que le nom le plus fréquemment utilisé sorte en premier lorsqu'il y a ambiguïté ; A. Oberto s'y emploie avec l'aide des astronomes. Le fichier de priorité de choix des noms a été révisé.
- Liens entre NED et Simbad: pas de réponse au message annonçant que l'outil est disponible. Discussion à mener à l'AAS.
- Specfind
Le papier a été accepté. Action de B. Vollmer et de F. Woelfel en janvier 2010: l'entrée du nouveau catalogue de cross-identification va doubler le nombre d'id radio (300.000).
- Raccord
Le nouveau raccord sera à reprendre après DJIN.

Logiciel et matériel

- L'affichage hiérarchie sera dans la prochaine version.
- Implantation récente des « parents virtuels ».
- Les liens Simbad/NED sont prêts.
- Le programme dicbuilder n'est pas loin d'être finalisé ; il est actuellement testé par M. Brouty. C'est un programme dans le même esprit qu'acrobuilder.
- Deux serveurs sur les trois de Simbad ont changé cet été. Ils sont plus puissants: 8 processeurs par machine (passage 64 bits). Intérêt: les machines virtuelles java peuvent avoir une puissance plus grande (ce qui a éliminé les problèmes de mémoire). Certaines réponses restent lentes. Il n'y a pas de gain sur le temps de réponse d'après F. Ochsenbein. La nouvelle version PostgreSQL est sensée être plus rapide que la 8.3 et devrait être implémentée.
- Cet été test sur 48 millions d'objets avec 160 millions d'entrées tables. Pas de problème sur les requêtes par objets mais ralentissement des requêtes par coordonnées.
- Le pourcentage de requêtes par objet est nettement supérieur au pourcentage de requêtes par coordonnées.
- F. Ochsenbein a mis au point une url en cours d'implantation dans les pages web de Simbad pour interroger les objets de VizieR par coordonnées. M. Wenger va implanter un lien vers le portail.
- Les identificateurs principaux: dans la page d'option, on peut choisir un ordre mais ce qui est mis là est également pris comme identificateur principal...

- Lien vers Simplay
A installer. Note: beaucoup de développements flash sont extrêmement sensibles à la version.
- Reprise complète du contenu et de la présentation de la documentation en ligne de Simbad. Ajouter la possibilité de naviguer entre les pages.
- Statistiques sur les derniers mois.
Suppression d'une requête hebdomadaire qui parasitait les statistiques.
160.000 requêtes par jour avec des pointes à 1.000.000.
Forte évolution des requêtes faites avec des applications et des programmes.
Action: identifier les requêtes web.
- Annotations
C'est une action développée dans le cadre des R&D d'EuroVO-AIDA. T. Boch a demandé à B. Gassmann de développer un panneau d'administration ; en pratique, ce service permet d'ajouter des commentaires (sous forme de texte libre) sur un objet donné, afin de compléter l'information disponible dans Simbad ou de signaler des erreurs éventuelles.
Ce développement va entrer dans sa phase de test et de validation technique prochainement.

VizieR

- Entrée des catalogues: le nombre augmente d'où une charge également en augmentation pour P. Vannier et F. Ochsenbein.
- Hardware
Changement de machines au fur et à mesure pour un passage en 64 bits, ce qui pose quelques problèmes, certains programmes devant tourner en version 32 bits. La reprise est donc progressive.
- G. Landais a installé une nouvelle version pour l'accès aux catalogues, avec un arbre de visualisation. Modification de la présentation ReadMe, accès à des données additionnelles... Une nouvelle version de VizieR est en cours de développement avec une partie historique gardant trace des requêtes précédentes.
- Un accès à VizieR par Topcat est maintenant possible. C'est très ergonomique. L'accès à toutes les tables d'un catalogue qui en comporte plusieurs ne marche pas avec Aladin mais marche avec Topcat.
- Statistiques
Les statistiques des pourcentages d'accès à VizieR via l'interface native et autres interfaces pourraient être mises en interne sur le wiki ?
- Problèmes identifiés pour le Registry. Il y a plusieurs points à voir et notamment les anciens UCD cone search. Tout n'est pas corrigé, il faut faire tourner le validateur (urgent).
- Problème au redémarrage, le cluster est fragile. Nécessité de faire des tests après chaque arrêt. Pas de problème d'homogénéité. Il faut remplacer les « mauvaises » machines et changer la machine qui sert à préparer les catalogues et sert aussi de machine de secours en prenant le relai chaque fois que le cluster ne fonctionne plus.
- Action de F. Ochsenbein: faire une présentation de la nouvelle version de VizieR en amphi, en l'annonçant à l'ensemble de l'Observatoire.

Aladin

- Nouvelle version depuis septembre 2009. Pas de souci majeur, juste quelques petits bugs résiduels.
- Remplacement du vieux serveur http (le 15 octobre). Suppression du vieux matériel et remise à jour des pages.
- Suite à la dernière réunion Aladin, 3 points ont été évoqués: la fiabilité, la simplification de l'interface et l'implantation du mode d'accès All Sky Healpix.
7 surveys sont accessibles en beta test:
 - IRIS 12 et 100
 - Fin du DSS couleur
 - A venir: Sloan ou DSS quick v
 Trois machines: le serveur Alaska et 2 machines de calcul (qui préparent les survey).
 Healpix local en cours d'évaluation par l'équipe de Planck.
 Adaptation de l'interface pour le mode allsky.
 Mise en place de la librairie java Healpix en discussion avec W. O'Mullane.
- Deux plugins Aladin sont en cours de développement avec l'équipe du LSIT ; extraction de spectre à partir de cube (bel outil).
Génération de SED à partir d'images: G. Mantelet s'en charge. Avec un lien avec le modèle caractérisation.
- Présentation d'Aladin à ADASS, au GIRES, et du plugin VOEvent à la réunion Interopérabilité de Garching.
Bons retours.
- Présentation à l'ASOV de Healpix par F. Bonnarel.
- Outreach: interrogation par arbre de catégorie d'objets puis interactions avec Massimo.
- Aladin sera utilisé pour le JWST et l'ESA l'utilise aussi pour visualiser les images solaires.
- Changement de machine couplé avec un changement de serveur ; reprise complète de l'astrométrie. Des bugs sortent sur les mosaïques. Aussi demandé de façon urgente: intégration de Glimpse. Il y aura une version test avant les vacances.
- XMM
Le souci rencontré après la transition au nouveau serveur images a été résolu.
- Beaucoup d'événements VO au mois de novembre
- Implantation de SIA2 est très avancée et sera à mettre en ligne. L'intérêt du serveur est de fournir les métadonnées.
- On fournit pour le plugin développé par G. Mantelet et C. Bot la caractérisation des observations: axes du SED.
- L. Cambresy a testé la possibilité de contrôler Aladin depuis un script IDL (feedback).
- Cluster
Utilisation massive du cluster et iRODS (50.000 tâches de 50 à 60 secondes chacune, c.à.d. plusieurs jours de calcul) dans le cadre d'un projet scientifique soutenu par AIDA.
Il y a un souci de dispatching en raison de l'utilisation massive, qui est en cours de correction. Un papier est en cours et peut-être un autre...
- R&D pour la détection de LSB: un papier est prêt et sera soumis en janvier. Des tests ont été faits pour l'utilisation sur le NGVS.

VO

- V. Meslard est en CDD de 6 mois sur le contrat européen AIDA sur les aspects sécurité, authentification,....

Il y a des échanges avec Astrogrid (mais leurs développements sont fortement dépendants d'outils extérieurs).

- Les wiki sont virtualisés sur une machine.
- Dans l'optique du Portail, V. Meslard pourrait épauler T. Keller pour l'installation (inclus dans AIDA).
- iRODS prototype: on est passé d'un prototype à une configuration stable ; il n'y a pas eu de problème samedi lors du transfert des machines.
- Grille
Réunion à Paris le 14 janvier prochain pour faire le point. Les italiens (INAF) sont occupés par le montage d'un projet européen.
- Réunion Grille-France d'octobre à Lyon
Transition EGI (European Grid Infrastructure) en cours de mise en place en France ; il devrait y avoir de la main-d'œuvre et des budgets.
- Workshop data repositories SHAMAN (8 millions de financement sur 4 ans).
Le projet est intéressé par une confrontation avec les données astronomies: à voir.

