

Compte rendu de la réunion de suivi CDS du 19 août 2003

Présents: M. Allen, F. Bonnarel, L. Cambrésy, S. Derrière, P. Fernique, F. Genova, B. Vollmer

Excusés: P. Dubois, S. Lesteven, F. Ochsenbein, A. Schaaff, M. Wenger

Bibliographie/Contenu

- Le log ISO a été rentré dans SIMBAD (E. Davoust, A. Salama, F. Woelfel). Une innovation très importante a été développée pour cette occasion: l'interface SIMBAD va chercher les données dans Vizier. Cela ouvre la possibilité de rendre plus étroits les liens SIMBAD/Vizier, en donnant directement accès aux données originales des catalogues.
- B. Vollmer a commencé à travailler sur la cross-identification des catalogues radio. Il y a environ 400 catalogues radio... Le point de départ est d'étudier les cross-identification des catalogues disponibles dans une même fréquence, en tenant compte de leurs caractéristiques (résolution, sensibilité, bande de fréquence). Lorsqu'on tient compte de la résolution et de la sensibilité, on voit immédiatement apparaître des hiérarchies entre les objets, les relations entre les sources peuvent être du type 'parent', 'enfant', 'frère/soeur'. Plusieurs problèmes:
 - il faut au moins trois fréquences pour décider si une source vue dans différents catalogues est bien la même (pente du spectre);
 - problèmes de confusion, en particulier dans les amas de galaxies;
 - on peut avoir besoin de partir de l'image NVSS quand elle existe.Gestion de la hiérarchie:
 - on peut ajouter des colonnes dans le catalogue pour les cross-identifications ou associations hiérarchiques;
 - faut-il ajouter les hiérarchies dans SIMBAD? – question soulevée depuis longtemps et très complexe (mise en oeuvre logiciel et surtout mise en place du contenu!). Des développements plus simples à mettre en oeuvre seraient utiles dans ce contexte;
 - types d'objets /longueurs d'ondes multiples (par exemple en donnant accès au type d'objet par défaut de tous les acronymes de l'objet);
 - accès à la liste des objets voisins ('customisée' selon les caractéristiques des objets).

Logiciel SIMBAD

- Voir ci-dessus la discussion sur les nouvelles fonctionnalités.
- Commentaires objet: à relancer DE TOUTE URGENCE. Il faut qu'ils soient mis en fonction en septembre. La version en test de la mise à jour contient d'autres modifications à tester d'urgence pour ne pas encore retarder la mise en service des commentaires objets.

Catalogues/VizieR

- Le stage de S. Nicaisse se termine. Le serveur devrait être mis en exploitation mi-septembre après validation par F. Ochsenbein. S. Nicaisse a rédigé la documentation (installation d'un cluster avec CLIC, utilisation du dispatcher. Une ré-installation complète a été réalisée en juillet afin de valider la procédure d'installation du cluster (pourra être utile pour le stage sur le serveur Aladin).

Aladin

- Une nouvelle version du logiciel intégrant les développements réalisés pour la démo AVO de juillet devrait être rendue publique en septembre ou octobre. A noter: le besoin exprimé par le projet GOODS (accès à l'arbre hiérarchique des données – voir ci-dessous); 'mosaïquage'; tracés de champ de vue d'instruments. La nouvelle version intégrera la bibliothèque JavaGLU, qui permet de prendre en compte beaucoup plus facilement les contraintes sur les différents ensembles d'images, qui ne sont plus globalisées (par exemple: pour tel ensemble d'image, passer par la construction d'une mosaïque), et de faciliter l'inclusion d'accès à de nouveaux serveurs externes. L'inclusion de cette bibliothèque va nécessiter une série de tests approfondis. Des tests de l'outil pour 'mosaïquer' les images, développé en juillet, devraient être réalisés par P. Guillout, la fonction étant demandée aussi par XMM (l'outil utilise le programme SWARP d'E. Bertin pour le ré-échantillonnage, et on peut prévoir une version où le fond est retranché et une où il est laissé tel quel, au choix selon les données). M 33 est un très bon objet de test. Le passage des paramètres doit être simplifié. Deux problèmes restent à régler côté serveur. Il reste à décider si la nouvelle version sera rendue publique en une seule fois ou en deux étapes, la première permettant d'être aussi proche que possible de la date de distribution des données GOODS. P. Fernique fournira une liste des nouvelles fonctions avec une indication de l'étape où elles seront mises en fonction s'il y a finalement deux distributions successives du logiciel.
- Données 2MASS: les dernières données reçues sont vides... D. Egret était chargé des contacts avec le projet, à reprendre par L. Cambrésy.
- Transfert de données massives (pour les 2.4 Tb du premier data release du SDSS): une nouvelle étude de transfert en ligne est en cours avec l'aide de RENATER (F.X. Andreu). Fort goulot d'étranglement aux USA (mais il doit être possible de le contourner puisque de grandes masses de données sont transférées entre Fermilab et le CERN), puis dans une moindre mesure au niveau d'OSIRIS. Ce dernier va devenir critique quand le problème sera résolu aux USA, et il devrait perdurer plusieurs mois. A discuter avec Renater.
- Supercosmos (500 Gb): un transfert de données en ligne est également envisagé. On arrive à des taux de transfert de 100-150 Ko/s avec Cambridge.
- Relancer DPOSS – qui va être le dernier catalogue manquant...
- Stockage: il faut acheter rapidement 3 To. Il est décidé d'acheter cette fois des RAID IDE, et d'attendre pour acheter du 3ware RAID (recommandé par SDSS et Terapix parce qu'il est beaucoup plus rapide pour le transfert de données vers le CPU).

AVO/projets/standards

- La démo AVO de Sydney a été un succès. Elle était beaucoup moins strictement organisée que celle de janvier, car elle comprenait beaucoup plus de fonctions et n'était pas orientée vers un objectif unique bien défini. Cela a conduit à des démos plus diverses, orientées au cas par cas selon les souhaits et l'expérience du visiteur, qui ressemblaient beaucoup plus aux démos habituelles du CDS (la plupart ont été courtes car elles ont eu lieu pendant la pause café, mais certaines ont duré plus de 20 minutes). Logiciel stable (un problème avec le serveur de la base de données image –l'entrée des données 2MASS a fait ressortir un bug, corrigé par F. Bonnarel le 19 juillet). Très bonne connexion sur le réseau scientifique australien. Le temps de réponse à la première requête à la base de données est beaucoup trop long (il est encore plus long lorsqu'on donne accès à toute la base de données – en cours de correction). L'accès aux données locales fournies par A. Richards a été démontré et constitue une fonctionnalité importante: accès à toute donnée disponible par une URL. Certaines démos sont allées jusqu'au niveau du développement d'A. Micol, par exemple une démo sur l'accès aux données radio pour P. Diamond. Quelques démos importantes politiquement (employé au NASA Headquarters, directeur du NOAO). Pour info: la démo a été faite sous Windows (problème de la fenêtre propriétés sous Redhat 9, difficulté pour utiliser une visu externe), mais les parties sur l'accès à des données locales et les arbres étaient sous Linux. Après l'UAI, M. Allen a présenté la démo à Mont Stromlo (Canberra) et à l'Université de Melbourne.
- Plus spécifiquement, démonstration de VOPlot, avec pour l'essentiel des appréciations positives. La fonction est utile mais il faut définir son évolution: pour le moment l'outil n'est pas assez puissant pour la plupart des applications astronomiques, et il y a plusieurs concurrents (Mirage, TOPCAT). Les points de vue sur l'évolution possible sont pour l'instant différents (il reste en particulier à transformer le logiciel en applet compatible Java 1.1: Action F. Genova vers A. Kembhavi). Visite des développeurs indiens à l'occasion d'ADASS, calendrier à fixer (avant ou après la réunion) au retour de F. Ochsenbein.
- Autres démos: AstroGrid (applications de visualisation utilisant Globus, visualisation de cubes de données incluant histogrammes, contrastes et rotations. Implémentation difficile. CVO: données et catalogues sous une même interface, utilisation d'un modèle de données simple, data mining dans les données. NVO: inscription et envoi de requêtes à un registry.
- F. Genova participera à la réunion NVO les 3 et 4 septembre à Victoria. Point à faire par téléphone sur le statut des UCD avant la réunion (1^{er} ou 2 septembre). Contacter R. Williams (UCD/S. Derrière et F. Ochsenbein) et W. O'Mullane (registry-GLU/P. Fernique) d'ici là. C'est aussi une occasion de discuter du CFHTLS et des collaborations éventuelles CDS/VO canadien avec D. Schade (CADC).
- Contact pris avec le projet GOODS (via B. Hanisch), intéressé à utiliser l'arbre hiérarchique pour donner accès aux données pour leur distribution publique (fin août). L'arbre hiérarchique de données ne sera pas inclus dans la version publique d'Aladin à cette date. Nous pouvons les aider à construire un arbre de données 'à la IDHA' et leur suggérer d'utiliser la version de démo AVO en attendant la nouvelle version d'Aladin. Affaire suivie par M. Allen.

- MDA: les moyens obtenus sont notifiés. Le CDS bénéficie de deux ans de CDD qui peuvent être attribués en plusieurs fois. Le LORIA a obtenu un post doc INRIA. Réunion de démarrage mi-septembre.
- IDHA: l'interface MARSIAA pour Aladin commence à être mise en route. Stage J.-F. Nicolas: il sera présenté le premier septembre. Réalisation d'un outil permettant de générer du XML à partir du diagramme UML d'IDHA. Sources en cours d'homogénéisation par A. Schaaff. Elles seront mises à disposition sur le Twiki, où l'exécutable est déjà téléchargeable. Un Web Service a été mis en place pour Aladin.
- Stagiaire UTBM: 6 mois à partir de septembre. Construction d'un prototype de calcul distribué sur le serveur image. Achat de matériel à prévoir rapidement (2PC). Echanges techniques à prévoir avec les travaux réalisés pour VizieR (stage de S. Nicaïsse).
- UCD: intégrer le gros travail réalisé par A. Preite Martinez sur les mots clés FITS, les VOX, les remarques d'A. Richards sur la radio, le travail avec A. Szalay sur SDSS. La nouvelle structure des UCD est discutée par le Steering Committee. Problème majeur: l'organisation des données photométriques, sur laquelle il n'y a pas d'accord.
- Web Services: développement d'une librairie SOAP 'light' – WSEA – Web Service Easy Access. C'est une librairie Java d'une cinquantaine de ko qui permet de dialoguer avec un serveur SOAP dans une application Java en faisant l'économie d'une librairie comme AXIS (dont la taille n'est pas la même). Un autre WS est en cours de mise en place à la demande de C. Arviset (récupération d'informations diverses dans VizieR). Etude de l'outil open source Jboss 3.0 pour remplacer éventuellement Tomcat4 mais pas de gain en performances.
- DENIS: réunion le 30 juin. Les images seront-elles fournies par Paris ou par Lyon? A trancher par le projet (Lyon n'a semble-t-il pas toutes les images) (Contacter le projet: L. Cambrésy).
- Planck: L. Cambrésy devrait être coopté au groupe de travail sur le VO.

Informations politiques et administratives

- Arrivée de L. Arbousse (assistante de la directrice du CDS) le premier septembre. Elle partagera le bureau de J. Marcout.
- Le Twiki du CDS est installé, utilisé par les stagiaires et par quelques personnes (par ex. il y a une rubrique UCD).
- Demandes européennes : les deux demandes envoyées au printemps ne seront pas financées. La demande principale (VO-INT), présentée au même appel d'offre 'Infrastructures' que OPTICON et Radionet, est arrivée largement au-dessus de la barre. Le projet sera re-présenté au prochain appel d'offre du même type, prévu courant 2004. Les détails de la stratégie sont en cours de discussion.

Colloques et information utilisateurs

- ADASS:
 - tutorial: en bonne position pour Web Services, contacter D. Tody pour inclure l'arbre hiérarchique dans le tutorial DAL (F. Bonnarel);
 - résumés M. Allen (accès aux données), S. Derrière (UCD), Aladin, A. Schaaff (serveur de catalogues).

