

Compte-rendu de la réunion de suivi du CDS du 13 novembre 2006

Présents: M. Allen, L. Arbousse, T. Boch, F. Bonnarel, L. Cambresy, S. Derriere, P. Dubois, P. Fernique, F. Genova, S. Lesteven, C. Loup, F. Ochsenbein, B. Vollmer, M. Wenger

Excusés: A. Schaaff

Informations générales

- Le projet européen EuroVO-DCA a commencé le 1^{er} septembre dernier (pour une durée de 28 mois). Une réunion d'information d'1h-1h30 est à prévoir. Le site web est le suivant :
<http://cds.u-strasbg.fr/twikiDCA/bin/view/EuroVODCA/WebHome>
- Réunion générale ASOV qui se tiendra à l'IAP du 28 au 30 novembre prochains. Le programme est disponible à l'adresse suivante :
<http://www.france-ov.org/twiki/bin/view/ASOVFrance/Reunion2006>
- Tutoriel Euro-VO en juin 2007: à l'ESAC, à Villafranca.
- Février 2007 (en fait mars 2007): une réunion sur la spectroscopie sera organisée dans le cadre EuroVO-DCA (WP3: Support to implementation by the data centres) par l'ESA.
- Ray Plante a présenté sa candidature pour venir en tant « qu'invité », au mois de mai 2007. Il est le nouveau responsable du groupe Registry de l'IVOA.
Action: établir une liste des sujets à discuter.
- Nikolaj N. Samus, responsable du catalogue étoiles variables à Moscou, souhaiterait également venir un mois.
- Un poste NOEMI « Assistant de bibliothèque et de documentation » niveau AI est ouvert au CDS.
- Réunion européenne IST 2006 à Helsinki du 21 au 23 novembre ; participation: F. Genova, M. Depretz.
http://europa.eu.int/information_society/istevent/2006/index_en.htm
- Réunion de présentation d'Aladin le 20 novembre à 10h.
- Posters pour Paristic du 22 au 24 novembre 2006
Il y a 3 posters à faire: 1 sur les aspects texte, 1 sur les aspects image et 1 sur les aspects workflow et calcul distribué. M. Louys présentera un exposé sur les aspects Workflow.
- Réunion Workflow OV France le 21 décembre à Paris.
- Réunion Grilles OV France le 9 janvier 2007 à l'ENS-Lyon.
- Proposer à R. Ibata un séminaire reprenant l'exposé « the many faces of SIMBAD » présenté à ADASS.

Bibliographie

- Passage de la mise à jour sur Simbad4 depuis le 2 octobre dernier. Une liste de problèmes a été fournie à A. Oberto qui en a déjà réglé une bonne partie. Des améliorations sont à envisager (définitions des formats de sorties, ...). D'autres difficultés, pas directement liées au logiciel de mise à jour mais à l'environnement (statistiques, créations de codes journaux, ...), sont à régler par M. Wenger. L'équipe bibliographie remercie A. Oberto pour sa disponibilité.

- Il y a eu 2 réunions du groupe Bibliographie dernièrement:
 - o Le 10/10: un compte-rendu des réunions LISA V, réunion de l'INIST, Ecole de Goutelas et UAI a été fait et se trouve <http://cds.u-strasbg.fr/twiki/bin/view/Reunions/ReunionsBiblio2006>.
Après un tour de table, tout est à peu près à jour mais certains journaux ne sont plus traités (AstL, Arep, Msngr, Obs, PABei, PASA, PNAOJ, PPMtO).
Action: faire un choix parmi les journaux pour décider de ceux qu'il faut continuer à traiter.
 - o Le 10/11: réunion Simbad4 avec A. Oberto.
- G. Jasniewicz a été très pris par GAIA mais A. Oberto lui a fait un nouveau mode d'emploi et il s'est remis au traitement des messages sur le contenu de Simbad. Les messages des utilisateurs sont traités à Strasbourg. Restent ceux de B. Skiff.
- Les réunions =g= ont démarré en mai. La difficulté est de trier les listes à entrer dans SIMBAD ; des niveaux de priorité ont été établis:
 - o =g1= (haute priorité): au départ, ils représentaient les $\frac{3}{4}$, ce qui était trop, et, après trois réunions ils représentaient la moitié de ce qu'il y avait à traiter.
 - o =g2= : $\frac{1}{4}$ au départ et la moitié au bout de trois réunions.

Il y a actuellement 133 =g1= et 33 =g2=.

On est passé de 120 =C= à 77 et de 470 =g= à 340. Entre 10 et 15 sont traités chaque semaine. Cela a permis de soulever plusieurs questions dont les magnitudes du Sloan et les données fondamentales par thème et non par ordre chronologique.

Toute l'information sur les réunions est sur le Twiki CDS.

- F. Durret a proposé de réfléchir à une stratégie pour la hiérarchie dans les amas qu'elle connaît ; elle s'y mettra dès la fin du mois de novembre.
- Acronymes: réunion F. Genova/B. Vollmer/M. Wenger pour le nouveau logiciel. B. Vollmer travaille encore sous Simbad3 et M. Wenger traduit les nouveaux acronymes sous Simbad4. Pour les modifications: les transmettre pour le moment à M. Wenger.
- Mise à jour Simbad
B. Skiff envoie beaucoup de remarques: faut-il lui donner des clés pour générer les commandes de mise à jour ? Imaginer une solution intermédiaire ? Le problème serait alors la validation. Les corrections de position peuvent être faites de façon semi-automatique. Il faudrait inviter B. Skiff et lui proposer de venir 15 jours-1 mois.
- Contacter E. Davoust pour le travail sur les données de 2MASS. L. Cambresy le contactera cette semaine car il était prévu qu'il vienne 15 jours fin 2006. L'objectif est d'ajouter les magnitudes 2MASS sur des objets existants.
- A&A a changé de machine (PC fourni par EPS).
Depuis cet été, nous n'avons plus d'accès sur le serveur de l'AAS et le problème n'est toujours pas réglé.
La nouvelle machine astrobib est en cours d'installation ; toutes les nouvelles données bibliographiques vont aller sur cette nouvelle machine (qui est un serveur SUN).

Contenu

- Catalogues Radio
B. Vollmer va reprendre contact avec F. Woelfel. Il va faire très prochainement un raccord sur toutes les sources radio pour corriger les distances fausses. Le logiciel SPECFIND incluant les développements de B. Gassmann va être installé chez lui et il va tester plus de catalogues. Le logiciel SPECFIND a été repris par les russes qui font une base de données chez eux.

Observation de 250 sources (GPS) avec le télescope de 100m d'Effelsberg. Les données ont été réduites et il reste maintenant à rédiger un article.

- Reconnaissance de noms d'objets à partir d'un fichier pdf (travail de C. Bonnin): ça marche globalement bien, il reste la phase de validation du logiciel de reconnaissance.

Logiciel

- Simbad4: état d'avancement de Simbad au 13 novembre 2006
Simbad4 est utilisé pour la mise à jour depuis le début du mois d'octobre ; cela a permis de régler des bugs en local.
6 semaines pendant lesquelles Simbad3 sera resté « figé » puis basculement des utilisateurs vers Simbad4 avant la fin novembre.
Il reste encore beaucoup d'utilisateurs qui font des demandes de compte. Ce sont probablement des demandes envoyées de façon automatique. Attendre et évaluer une relance éventuelle.
Aladin: on ne peut pas vérifier, il faut créer un formulaire Simbad4 pour les bibliographes.
Règlement de problèmes et bugs qui arrivent pour la plupart de la mise à jour.
L'accès autour du RAID: la documentation est presque terminée et sera mise en ligne demain. Il reste des petites modifications à faire dans les URL puis il faudra regarder le switch des machines (Dictionnaire flu-redirections-Apache, firewall).
Les accès clients-serveurs se feront de façon transparente et la redirection sera automatique.

Catalogues/VizieR

- Statistiques
Il y a actuellement environ 100.000 interrogations par jour et cela va même jusqu'à 300.000.
- Matériel
La mise en place des clusters (Aladin et serveur catalogue) a bien avancé.
Au total, il y a 4 machines pour les catalogues et autant pour le calcul distribué.
- Miroirs
G. Landais a terminé son travail pour que VizieR tourne sous PostgreSQL. Il faudrait faire des comparaisons plus précises Sybase/PostgreSQL. On garde les 2 implémentations différentes ; le clonage se fait à partir de la machine Sybase. S'il y a un nouveau développement, ce devrait être transparent. Il peut y avoir des catalogues sous Sybase et sous PostgreSQL: à tester.
Discussion avec Alberto Accomazzi (ADS) pour l'installation de la nouvelle machine au CfA avec le même serveur PostgreSQL pour VizieR et SIMBAD. Le CfA a fait des changements sans nous prévenir, qui ont compliqué l'exploitation du miroir.
- Catalogues
P. Vannier et M. Brouty ont rattrapé une partie du retard et le travail avance bien.
Entrée des données de gros catalogues tels NOMAD (fusion USNO et 2MASS).
Arrivée de nouveaux catalogues comme :
 - o GSCII.3. avec 1 milliard de sources, publié avec des noms propres mais il ne faut pas les diffuser ; B. Skiff s'en occupe.
 - o SDSS 5
 - o MAPS: à reprendre pour l'avoir au moment de l'AAS.
- Interface utilisateur: en cours de discussion.

Faire une proposition de stage pour 3-4 mois (+ l'été) à un étudiant ingénieur sur une méthode pour améliorer l'interactivité des interfaces, à l'aide d'une méthode plus dynamique (AJAX). Le sujet est en cours d'élaboration et sera diffusé dans les « réseaux » d'A. Schaaff. Il pourrait être repris par VO-TECH...

- Pour mémoire: passage aux UCD1+ en août 2006.

Aladin/Images

- Lancement au mois de décembre d'Aladin dans APT: le résultat de la collaboration est très positif.
Il y a eu de nombreux intervenants au CDS et au STScI. Cette nouvelle utilisation augmente fortement la pression opérationnelle.
La diffusion d'une Note IVOA sur les Fields of View a pris du retard à cause d'un besoin nouveau exprimé par l'ESO: séparer géométrie de rendering, ce qui a pris 2 mois de plus (discussions en cours avec M. Dolensky et J.C. Malapert).
- Lancement des macros simplifié afin de rendre plus facile les traitements répétitifs sur Aladin. C'est particulièrement utile pour M. Brouty.
- Action: faire un tutoriel des nouvelles fonctionnalités pour le personnel de l'Observatoire (le 20 novembre au matin).
- Plugins Aladin: en réponse aux demandes du code source Aladin, augmentation des possibilités de plugin: on peut installer sa propre extension. P. Fernique finit les pages ; cela a été envoyé aux 2 équipes qui avaient demandé le code source et a obtenu un retour favorable. On développe un peu de java, on en fait une classe et on le met dans un répertoire. On pourra faire de la publicité pour faire une bibliothèque de plugin Aladin. Il faudrait faire une documentation générale « customisation d'Aladin ».
- Support du chinois
Amélioration lente: Chenzhou Cui s'en occupe. On aura ainsi des versions d'Aladin en français, anglais et chinois.
- CADC
Formulaire d'accès aux archives en phase de finalisation. Il y aura un nouveau bouton CADC avec 3 archives (CFHT, JCMT, HST).
- Evolution du logiciel: le code a été modifié pour permettre d'accéder à des images jusqu'à 2 gigas ; ça marche aussi pour les gros cubes de données.
La possibilité d'extraire des catalogues des images avec Sextractor (Terapix) sera en ligne (les calculs sont effectués sur le cluster).
L'interface est peaufinée pour simplifier la vie de l'utilisateur: gestion plus facile (ex: roulette souris).
Plastic presque OK.
Client: en-têtes WCS dans du jpeg (VO-TECH): ça marche, les images de Sloan jpeg sont compatibles.
- Concurrence toujours vive de DS9.
- Documentation Aladin. Il faut mettre à jour le manuel utilisateurs.
Action: en reparler à O. Bienaymé pour le faire en janvier.
- Prochaine release officielle: ce sera en janvier, V4.0 (compatible UCD1+). Il y aura une grosse amélioration sur les volumes de données manipulables.
- Projet Image: aspects OV
 - o Standards VO: description du FoV. Note retardée par demande ESO. A finir dans les semaines qui viennent.

- *Characterisation*: présentation à peu près définitive à Moscou. On commence à avoir des implémentations. I. Chilingarian et Ivan Zolotukhin ont présenté un poster à ADASS ; derrière, il y a un service de données du télescope de 6 mètres russe. Serveur Aladin: version visible dans DataScope qui renvoie *characterisation* en VOTable. B. Gassmann a sorti la partie *characterisation* du Registry (indépendamment du service catalogue): à faire voir aux co-auteurs (A. Richard et A. Micol). Le document est passé en working draft officiel il y a 3 semaines. Baroud d'honneur d'A. Rots, n'a pas été soutenu par J. McDowell. STC repris à bas niveau.
- DAL: l'extension dans SIA/SSA n'a pas été reprise pour le moment mais le sera dans la prochaine version. Upgrade à faire. F. Bonnarel est co-auteur de SSA au nom des extensions non encore intégrées.
- Serveur Aladin: matériel
Le disque tampon mis en place avant l'été est fiable.
Il reste une nouvelle machine à installer ainsi qu'un nouveau RAID. Un PC a été commandé pour 2007 (pour découpler le frontal SUN et les tâches en arrière plan).
- Serveur Aladin: logiciel
Aladin client peut gérer toutes les plaques de Schmidt, ce qui permettra d'éliminer les vieilles parties du logiciel ; les plaques MAMA ont été rentrées dans le nouveau logiciel cet été.
Il reste une dernière chose à faire: la compatibilité au niveau jpeg (intégrer au niveau serveur ce que P. Fernique a intégré au niveau client). Tout le ciel Sud MAMA est disponible dans la base test. F. Bonnarel n'utilisera plus Sybase.
- Passage du serveur Aladin en SIA extension (cela fonctionne bien au Canada) à faire rapidement: pour une uniformisation des différents modes d'accès à un dataset.
- CFHTLS
Réunion des utilisateurs français lundi et mardi derniers: c'est la première depuis mai 2005.
Aspect gestion de données: J.-J. Kavelaars (CADC) a réalisé et montré un serveur SIA extension (données structurées) et F. Bonnarel a fait une démonstration de l'aspect grandes images.
- LSIIT
 - Un stage d'approfondissement du travail de M. Petremand a eu lieu cet été avec M. Wolff sur la détection de LSB. Les résultats seront analysés dans une quinzaine de jours et il sera décidé d'une publication ou non (visite de W. Van Driel du 27 au 29 novembre). Il faudrait proposer à M. Wolff des vacances puisqu'il continue de travailler sur le projet.
C. Collet voudrait reprendre les choses en amont en utilisant des techniques différentes avec B. Gray.
 - Un nouveau maître de conférences, V. Mazet, doit reprendre les affaires de F. Flitti. Il ne connaît pas les techniques markoviennes mais connaît bien les techniques bayésiennes.
 - Thèse de M. Petremand le 20 novembre prochain sur les techniques de réduction de données hyperspectrales.
Il faut noter les difficultés inhérentes aux thèses interdisciplinaires.
- Atelier OV France organisé par E. Slezak sur la mise en ligne des données en 3D dans le VO (M. Louys, F. Bonnarel, C. Collet, A. Jalobeanu): cela a permis de montrer les développements cubes qui intéressent Marseille et Le Sidaner. A clarifier: le rôle de P. Prugniel dans OV 3D.

VO TECH / VO / MDA

- Du 4 au 7 septembre: réunion projet VO-TECH à Strasbourg qui a consisté à passer en revue les développements faits pour le cycle 3 et planifier ceux du cycle 4. Les présentations sont en ligne sur le Twiki VO-TECH
<http://eurovotech.org/news.htm>
Prévoir une réunion interne pour faire le point avant Noël.
- Note sur Plastic: tarde à être publiée (dans le cadre d'IVOA).
- Mise en route du travail de C. Pestel sur le projet VO-TECH, en parallèle du travail (finalisation des nouveaux clusters Calcul et Données) suite à son stage. Peu de maintenance technique pour les nouveaux clusters.
Réunion de mise en production prévue cette semaine pour les clusters de Calcul et de données. Deux nouvelles machines sont en commande pour VizierR.
- DS5: Intelligent Resource Discovery
 - o Librairie séparée pour *caracterisation* (remise en forme des outils). Il s'agit d'outils pour identifier les ressources intéressantes et avoir des tables homogénéisées.
B. Gassmann procède également à la diffusion de la documentation associée.
S. Derriere finit les tests sur les dernières versions des outils.
La suite du travail porte sur la mise en œuvre de la nouvelle version du Registry. Le groupe de travail registry va pratiquer des tests de la nouvelle version (passage au 1^{er} janvier) ; c'est à ajouter également dans la liste des travaux VO-TECH. Garder trace de ce qui est fait (les tests ainsi que l'implémentation) dans le DS5.
 - o Travail d'A. Richard et A. Preite Martinez: ontologie montrée à Prague, Poster ADASS, Note technique IVOA. Il y a des interactions avec différentes personnes pour le développement de certaines branches: L. Cambresy pour les régions diffuses, P. Padovani pour QSO ; il faut trouver des intervenants américains (contacter les gens de la liste UCD/Semantic pour trouver des volontaires).
1^{ère} application recherche des ressources dans le Registry ; essai de rendre l'outil plus robuste.
- Les réunions de travail des 2 groupes Workflow et Grilles sont en cours d'organisation.
Après 2 premières réunions consacrées en grande partie à des présentations, le groupe Workflow s'oriente maintenant vers une réunion de travail (dont 3-4 personnes du CDS) avec un nombre de participants plus restreint: 20 personnes maximum.
La réunion Grilles devrait permettre de développer les relations avec les collègues STIC de Lyon, Grenoble, Nice, etc...On devrait dégager au travers de cette première réunion un point de liaison avec le WP DCA Grid (WP5) piloté par l'INAF.
- MDA se termine le 21 décembre prochain.
Les pages Web publiques sont à finaliser si possible d'ici la réunion Paristic et au plus tard fin décembre pour la fin du projet.
- Souci: achèvement du travail AIDA. Le stage de B. Peret a permis d'avancer mais trop court pour terminer le travail à faire.
Des décisions sont à prendre en ce qui concerne le niveau de maintenance des développements réalisés dans le cadre MDA, en particulier AIDA.