

Compte-rendu de la réunion de suivi du CDS du 14 novembre 2005

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

Excusé: M. Allen, L. Cambresy, A. Schaaff

Informations générales

- Ecole d'Obernai: bonnes retombées; beaucoup de personnes de nombreux laboratoires sont impliquées ou souhaitent s'impliquer dans le VO. Création d'un groupe de travail « images » (responsable: E. Slezak).
- Les pannes électriques posent des problèmes. Une solution serait d'avoir deux alimentations.
Pour l'onduleur, il faudrait avoir une maintenance régulière pour faire les tests. Il faudrait également avoir une formation. Un « arrêt maintenance » serait envisageable tous les six mois et non pas tous les mois.
- Le CDD VizieR est en phase de recrutement: il s'agit de Gilles Landais, pour le 1^{er} janvier 2006.

Bibliographie

- G. Chassagnard a repris son 80%.
- Réunion bibliographie les 15 et 16 septembre derniers. Présentation de SIMBAD, ce qui a poussé l'installation des PC pour pouvoir procéder aux tests. Ce sera fait demain pour les parisiens. Une réunion de formation (sur Linux et KDE pour la configuration des PC) devra être organisée: par qui et quand?
Action à mener dès que possible.
- Réunion à organiser avec les bibliographes sur les flux/magnitudes avant la mise en route de SIMBAD 4: à discuter au préalable avec M. Wenger, A. Oberto, L. Cambresy et B. Vollmer.
- Pré cross-id de 2MASS: gros travail pour les objets brillants; action à mener avec L. Cambresy.
Le problème (depuis 2003) est que 4000 2MASSi ont été rentrés en référence ainsi que des 2MASS1: il va falloir tout « nettoyer ». Pour cela: il y a le catalogue NOMAD de Monet qui a toutes les correspondances 2MASS.
Action de M. Wenger: faire un essai pour voir si on peut aller jusqu'à 20 millions d'objets dans SIMBAD.
Chenzou Cui fait des tests d'entrée de l'USNO B dans des bases de données; il arrive à 180h.
- Démarrage de réunions régulières pour le tri des =g= (pour diminuer leur nombre et rationaliser les critères de sélection). Ces réunions ont lieu à peu près toutes les

semaines, avec C. Brunet, F. Woelfel, L. Cambresy, B. Vollmer, F. Ochsenbein, P. Dubois.

Une réunion autour des procédures d'entrée des =g= est également à organiser avant le début de l'année prochaine. C. Brunet a commencé à expliciter les procédures.

- GCN (circulaires d'informations sur les ray brusts).
Il y en a 1340 en 2005 d'où une forte augmentation qui crée des problèmes. A discuter avec S. Laloe. Il y a 25% d'erreurs sur les auteurs. Il faudrait améliorer le logiciel de détection des auteurs. Cela demande quand même une ½ journée de travail par semaine. Sinon, on peut être moins exhaustif qu'ailleurs sur ce point.
S. Lesteven reste le point de contact.
- S. Lesteven a récupéré un stage de 3 étudiants ESIAL pour la reconnaissance semi-automatique des noms d'objets. Suivi serré à faire.
- Demande INSU pour le renouvellement de la machine biblio: faite dans le cadre de la Jouvence-CDS.
- FUSE: C. Bot a avancé dans l'identification du catalogue FUSE avec SIMBAD. Il lui reste 40 cas à faire qui sont des cas critiques et qui seront traités en collaboration avec J.-M. Désert. C. Bot a fait tourner la dernière liste d'identifications dans raccord avec l'aide de C. Brunet. Il reste à sortir une colonne contenant les identifications SIMBAD dans le catalogue FUSE traité. Pour le moment, les objets du système solaire ont été enlevés, ainsi que les données du milieu intergalactique. A voir aussi avec F. Ochsenbein. Reste aussi à quantifier les objets trouvés par raccord, et ceux identifiés par examen au cas par cas.
- Acronymes
B. Vollmer crée environ 20 acronymes par semaine. F. Genova s'occupe de ceux qui sont urgents ou qu'il faut modifier ; les piles diminuent. On est à peu près à jour pour 2004 alors qu'il y avait une accumulation importante depuis mars 2004.

Contenu

- Radio:
 - o LSB:
Réunion avec B. Vollmer, M. Louys et M. Petremand.
Pour le moment, on a seulement traité 4 plaques sur une centaine. Nécessité d'organiser la visite de W. van Driel pour les 28-29 novembre 2005, en présence de C. Collet. Printemps 2006: prévoir un stage sur des données multi longueur d'ondes (ex: CFHTLS).

Catalogues

- Tout sera à jour d'ici la fin de l'année pour les catalogues en retard de 2004. Puis, M. Brouty passera aux =g= (faire en sorte que ce soit avant le départ pour congés maternité de C. Brunet tout en conservant des =f=). Depuis 1993, ce travail s'est fait en ajustant le travail au fur et à mesure des recrutements mais commence à relever d'un régime permanent. Il reste à évaluer le travail pour l'ajout des UCD.

- Matériel
Le PC mis en route pour le ftp est opérationnel. Pour le cluster d'accès à Vizier, il faut ajouter des time out dans le gestionnaire du cluster pour éviter les blocages (quand il y a 300 process en attente, le PC s'écroule et ne sait plus gérer).
- Arrivée de nouveaux gros catalogues:
 - NOMAD (+500 millions d'objets): cross-identification de grands surveys fait à l'USNO par D. Monet.
 - SDSS 3 (environ 200 millions d'objets): est en ligne.
 - SDSS 4: ne devrait pas tarder.
- Tendance à noter:
Le CDS reçoit des bibliothèques de spectres synthétiques ; les valeurs des paramètres sont stockées dans Vizier mais les spectres eux-mêmes ne sont pas dans les tables (il faut faire des interpolations à partir des spectres synthétiques). Ces bibliothèques sont de plus en plus utilisées.
- Réunions régulières du groupe catalogue le mardi pour résoudre les problèmes rencontrés et pour la mise au point progressive des différentes procédures.

Logiciel

- Simbad4: état d'avancement de Simbad au 14 novembre 2005.
La présentation est orientée vers ce qui reste à faire...

- 1- Interrogation de la base
- 2- Application de mise à jour
- 3- Applications annexes
- 4- Données
- 5- Déploiement
- 6- Présentation CDS

1) Interrogation de la base

Il reste à terminer le langage de création d'échantillonnages. Le logiciel est opérationnel, mais il reste à décrire la majorité des mots-clés qui permettront les requêtes sur toutes les données (y compris les mesures).

La sortie VOTable est en cours de réécriture, suite à un changement de stratégie: affichage d'une donnée principale unique pour chaque objet dans une première table et report des données suivantes dans des tables secondaires.

Cela va permettre une utilisation simple et cohérente dans les autres applications VOTable et une insertion directe dans TOPCAT. Reprise de SAVOT.

Il faut aussi revoir la définition des UCDs pour chaque donnée (modif de tables)

Le lien avec le service biblio est à compléter.

Il reste à intégrer le calcul complet des coordonnées (avec pm, rvel & plx) dans le package Astro et intégrer cela dans le calcul des coordonnées (à fournir par F. Ochsenbein).

Voir les définitions de flux et les fonctionnalités autour.

La gestion des plots en utilisant Aladin est à faire (interface et filtres Aladin à définir).

Quelques petits détails seront aussi à améliorer: optimisation des requêtes sur des longues listes, rendre transparent certaines requêtes par bibcode (sans car M par exemple).

L'accès par programme à la base est aussi à développer:

- accès de type ASU (utilisé par Duvert notamment)
- accès par Web Services à définir

2) Application de mise à jour

L'interface graphique est actuellement en phase de mise au point. Tests et corrections de bugs.

Acrobuilder est toujours en cours de développement. Refactoring terminé après un blocage...

Acronymes: l'application est en cours de développements.

A voir: les procédures de correction batch de bibcodes et d'identificateurs

3) Applications annexes

Accès batch à Simbad fonctionne, mais l'ergonomie des commandes pourra être revue en fonction des besoins.

Raccord est à tester, a priori simcli tourne.

Une interface simple vers info actuel est à mettre en place, ainsi que quelques autres appli externes (findXXX)

Va devenir une commande standard de Simbad.

simbib & simref: serveur annexe à écrire. Il s'appuiera sur l'API Simbad qui réalise ces fonctions.

adsbib (liste des nouvelles références récupérées par ADS tous les 10 jours) est à reprendre.

4) Les données

Une base à jour vient d'être créée début Novembre. Il n'y a quasiment pas de problème de données.

Une statistique complète sur les différences sera faite sur cette version.

Il reste à faire:

- l'exploitation des liens hiérarchiques et associatifs
- l'inclusion dans les requêtes de différentes tables annexes (DOI, bibcodes synonymes, etc ...)

A noter sur ce point le travail d'O. Malkov qui a repris des problèmes sur les galaxies qui étaient en suspens depuis le départ de Paturel.

5) Déploiement

Il est prêt, modulo une machine actuellement HS.

Il reste à compléter le développement des moniteurs des serveurs, mais un minimum vital est opérationnel.

Il reste encore à implanter de façon exhaustive les enregistrements de log dans les endroits stratégiques.

6) Calendrier

Cette semaine: nouvelle version de la base disponible sur le site web

Fin de l'année: toutes les fonctions disponibles. Web services ?

Début 2006: démarrage dès que tout roulera comme il faut, que la mise à jour sera opérationnelle (formation), et qu'il y aura du support garanti pendant une ou deux semaines après le démarrage.

Aladin

- Aladin Java
 - o Les développements portent sur:
 - l'extension de l'interface ExtApp en VOApp pour permettre la manipulation d'autres données que des VOTable (P. Fernique): terminé
 - développements pour les FoVs dans le cadre de la collaboration STSci/Aladin (P. Fernique, T. Boch et F. Bonnarel): terminés en phase de tests
 - Chargement des FoVs sous forme de fichiers XML
 - Manipulation des FoVs (rotation, sous-éléments...)
 - Gestion de la synchronisation APT ; Aladin via VOApp
 - Implantation de la génération d'images Mosaïques (P. Fernique): en test
 - Gestion des serveurs d'images interrogeables par champs rectangulaires sur le ciel (et non uniquement par cônes) (P. Fernique): terminé

- Conversion des images RGB en niveaux de gris (P. Fernique): terminé
- o Les collaborations sont en bonne voie:
 - ESO (N. Delmotte): l'intégration d'Aladin dans l'archive ESO est effective depuis l'été. Des développements sont encore en cours pour essayer de charger directement des previews et pas uniquement des cartes de champs
(voir http://archive.eso.org/eso/eso_archive_main.htm)
 - ESA/XSA (C. Arviset): l'intégration d'Aladin dans XSA est effective depuis octobre ; elle s'étendra aux autres archives (ISO,...) au fur et à mesure de leur release.
 - STScI/MAST (R. Thompson): l'intégration de MAST est terminée et réellement effective depuis fin octobre.
 - STScI/APT (T. Donalson): les développements Aladin pour APT sont terminés et T. Donalson est en train de tester le remplacement de VTT par Aladin. C'est en bonne voie.
 - TopCat (M. Taylor): intégration de l'interface ExtApp dans TopCat pour pouvoir faire interagir Aladin et TopCat. En attente de la décision de M. Taylor pour l'intégration du successeur de ExtApp (VOApp).
 - VOPlot: N. Urunkar (nouveau développeur) est d'accord pour intégrer VOApp.
- Serveur image
 - o Serveur intégration plaques de Schmidt: implémentation prototype. F. Bonnarel travaille à présent sur les scripts. Changement d'architecture (essai avec Postgres): F. Bonnarel va rédiger un document préliminaire d'ici fin décembre.
 - o Collaboration C. Collet et M. Louys: Gallics
Bugs de notre côté mais le problème a été résolu la semaine dernière suite à une discussion avec E. Slezak.
Résultat sur le cube de données du centre MSI (Michelson Spectro Interferometry) fourni par E. Emsellem.

OV / MDA / VO TECH

- Réunion ACI STIC à Bordeaux du 21 au 23 novembre prochains.
F. Genova présentera un papier oral d'1/2 heure et insistera sur l'intérêt du VO astronomique pour les STIC.
- MDA
 - o Visite de N. Hernandez qui passera sa thèse début décembre. Réunion onthologie. N. Hernandez n'aura pas le temps de valider tout le thesaurus ; quelle stratégie faut-il proposer ?
 - o Réunion calcul MDA du 17 octobre
Il faudrait définir avec S. Genaud un sujet de stage pour MDA ; il serait important d'avoir un stage sur l'utilisation d'EGEE (30 millions d'euros obtenus, et une seconde tranche demandée, pour le développement de couches d'accès à la grille).

- J.-J. Claudon a repris le côté implémentation mais problème de performance ; sa thèse est le 8 décembre prochain. A voir avec F. Flitti qui restera ATER après sa thèse.
- Data Model / DAL
 - Modèle Caractérisation:

F. Bonnarel pensait sortir une note après Madrid mais cela a traîné. J. MacDowell est prêt à donner son accord pour cette semaine ; les commentaires sont en cours de relecture.

Une discussion est en cours sur l'utilisation du modèle caractérisation par VizieR et pour les développements VO-TECH de B. Gassmann: version serialisation ASCII. Faire des sorties du serveur Aladin qui donneront la caractérisation.
 - SIA extension:

Discussion avec D. Durand (qui est là jusqu'à fin janvier) qui a mis en place la dernière version du script faite par F. Bonnarel. Téléconférence avec J. Kavelaars pour CFHTLS: possibilité d'extension au-delà, à toute l'archive CADC et en particulier CFH12K. Intégration possible sur les données du CGPS ? Une discussion est en cours entre F. Bonnarel et D. Durand.
- Discussion avec Marseille
- SITools: discussion entre C. Surace et F. Bonnarel. C. Surace a commencé à intégrer leurs bases de données dans SITools ; ils travaillent sur les sorties VO. Plusieurs bars SIA-SSA.
VO-3D: ils voudraient utiliser Aladin comme couche de recherche de données (R. Savalle).
- GRID: réunion du 20 octobre au CNES.
Il y a des applications de la grille de calcul qui paraissent très prometteuses.
Il faut également rediscuter avec F. Capello (qui a dépêché P. Marty à Obernai) de Grid 5000 qui est une grille expérimentale complètement reconfigurable et qui commence à être opérationnelle.
Il faut aussi contacter L. Fusco à l'ESA concernant l'utilisation réelle de la grille et l'inviter à un séminaire au CDS.
- Déplacement d'un à deux jours prévu d'A. Schaaff et de J.-J. Claudon en décembre à Nice pour une démo et des discussions AIDA.
- PADOUE: le projet est terminé
Rédaction d'un court rapport par A. Schaaff à inclure dans le document final (le CDS a participé en tant qu'expert).
- Visite de nos collègues de l'OV japonais (du 10 octobre) Yuji Shirasaki et Masahiro Tanaka. Travail d'implémentation de Skynode et intérêt pour les affaires workflow. Discussion avec A. Schaaff sur les aspects DAL. Discussion avec F. Bonnarel des relations avec SIA. L'interfaçage avec Aladin est en test. ADQL→SIA: ont cette couche, ce n'est pas la peine de la refaire.
- Visite de collègues de l'ESA: P. Osuna, A. Stebe et I. Ortiz

Objet de la visite: implantation d'un Skynode dans VizieR, pour une discussion concernant Registry, préparation d'une liste de commentaires-demands à l'intention du groupe VOQL dans la foulée.

Demos CDS, actions OV France

- OV France
1^{ère} journée workflow a priori constructive (beaucoup d'exposés et de questions), quelques tâches ont été définies et une téléconférence est prévue vers la mi-janvier.
Il reste à fixer une date pour le workshop Web Services (avant ou après le workshop OV France standards VO de la fin janvier).
Workflows: discussion sur la multiplicité des systèmes: faut-il en choisir un ou tester ce que peut faire chacun ? Pas de conclusion. Proposition: fournir des cas d'utilisation dans différents domaines spectraux, théorie.

- American Astronomical Society du 8 au 12 janvier 2006 à Washington
Il y aura un stand de démo pour le CDS. Il faut s'inscrire avant le 2 décembre. Suite à la mauvaise connexion de la dernière fois, nous aurons droit à une réduction sur la connexion (besoin d'un niveau où l'on a 5 adresses). Qui ira à Washington ? F. Genova, L. Cambresy, M. Wenger, S. Derriere ?