



Atelier SGSD-A 2005

Systeme de Gestion des Donnees
en Astroparticules

Lyon 2005

<http://sgsda.lal.in2p3.fr/>



La situation

- Nombreux efforts autour de la question du stockage et de l'accès aux données
 - Souvent dupliqués, parfois contradictoires
 - Cas de PIOLib dans Planck
 - Utilisation de solutions « partiellement » appropriées
 - Objectivity dans BaBar, Versant dans Planck
- Différence d'approche entre les astronomes et la physique « accélérateur »
 - Voir la présentation « [Observatoire Virtuel](#) »
 - Situation en Astroparticules
 - Beaucoup plus de données qu'en astronomie; beaucoup moins que dans LHC



Les débuts

- Quelques représentants de Eros, Auger, Planck, Snls - physiciens et ingénieurs du LAL, CEA, Cdf & APC
 - Expériences confrontées à de grandes quantités de données (< 100 To)
- Plusieurs réunions en 2004 destinées à confronter les approches et les expériences



SGSD-A

- Réponse à PIAP (*Programme Interdisciplinaire Astroparticules*)
- Objectifs
 - Confronter différentes approches et solutions à des cas concrets en dehors des contraintes d'une expérience
- Financement pour un atelier en 2004
 - En attente d'une réponse pour 2005
 - Intérêt du Centre de Calcul de contribuer à l'étude
 - 2 expériences « virtuelles » (satellite et observatoire)
 - Volumes réalistes de données



Atelier Lyon 2005

- 2 jours de réunion au CC IN2P3
- 15 participants
 - LAL, IAS, APC, CEA, CERN, CCIN2P3
- 5 thèmes
 - Les formats de fichier
 - Les bases de données
 - Les serveurs de données
 - Pool
 - L'Observatoire Virtuel



Les formats de fichiers

- Formats généraux
 - Root
 - FITS
 - HDF5
- Les formats « expériences »
 - PIOLib
 - Sophya



Les bases de données

- Les bases de données relationnelles
 - Concepts, intérêts
 - Oracle, MySQL
- Les bases de données objets
 - Points forts, faiblesses
 - Objectivity, Versant
- Les ponts Objets Relationnels
 - Utiliser l'approche objets avec une base de données relationnelles



Les serveurs de données

- Caractéristiques des systèmes de stockage de masse
 - Le robot de cartouches HPSS
- Serveurs de fichiers distribués
 - SRB
- Serveurs de données
 - Xrootd



Pool

- La solution de stockage, de gestion et d'accès aux données du LHC
- Des contraintes et des solutions à l'échelle des expériences LHC
- Choix de la grille



Observatoire Virtuel

- La démarche de l'Observatoire Virtuel
 - Mise en commun des observations réalisées par la communauté
 - Effort de standardisation des formats et des méthodes
- Le cas de l'expérience EROS
 - Mise à la disposition de la communauté d'une incroyablement riche collection d'images du ciel et de résultats scientifiques (courbes de lumière)



Conclusions

- Dépend en partie de la réponse à la proposition PIAP pour 2005
- Accord du CC IN2P3 pour participer à l'effort matériel
 - Machine, espace disque, accès aux logiciels déjà disponible (Oracle)
- Définir concrètement un (ou plusieurs) modèle expérimental, plusieurs cas de productions, et différents solutions logiciels
- Evaluer ces solutions pour des volumes réalistes de données dans une simulation raisonnable de production (nombre important de processus)