

Computing in High Energy and Nuclear Physics

Résumé des sessions «Distributed Event Production and Processing » et «Distributed Data Analysis »: « la grille vue par les expériences » par Vincent Garonne

Séminaire interne LAL -16/03/2006



Aperçu général

- Le « Service Challenge »
- La production de données
- L'analyse des données
- Conclusions



« *Service challenge* »

- *Du point de vue de LCG:*
 - Renforcement du middleware
 - Test de nouveaux services (e.g. données)
 - Grande participation des sites (Tiers1 + 20 tiers2)

- Du point de vue des expériences
 - Permis de tester les modèles de calcul, les environnement de production

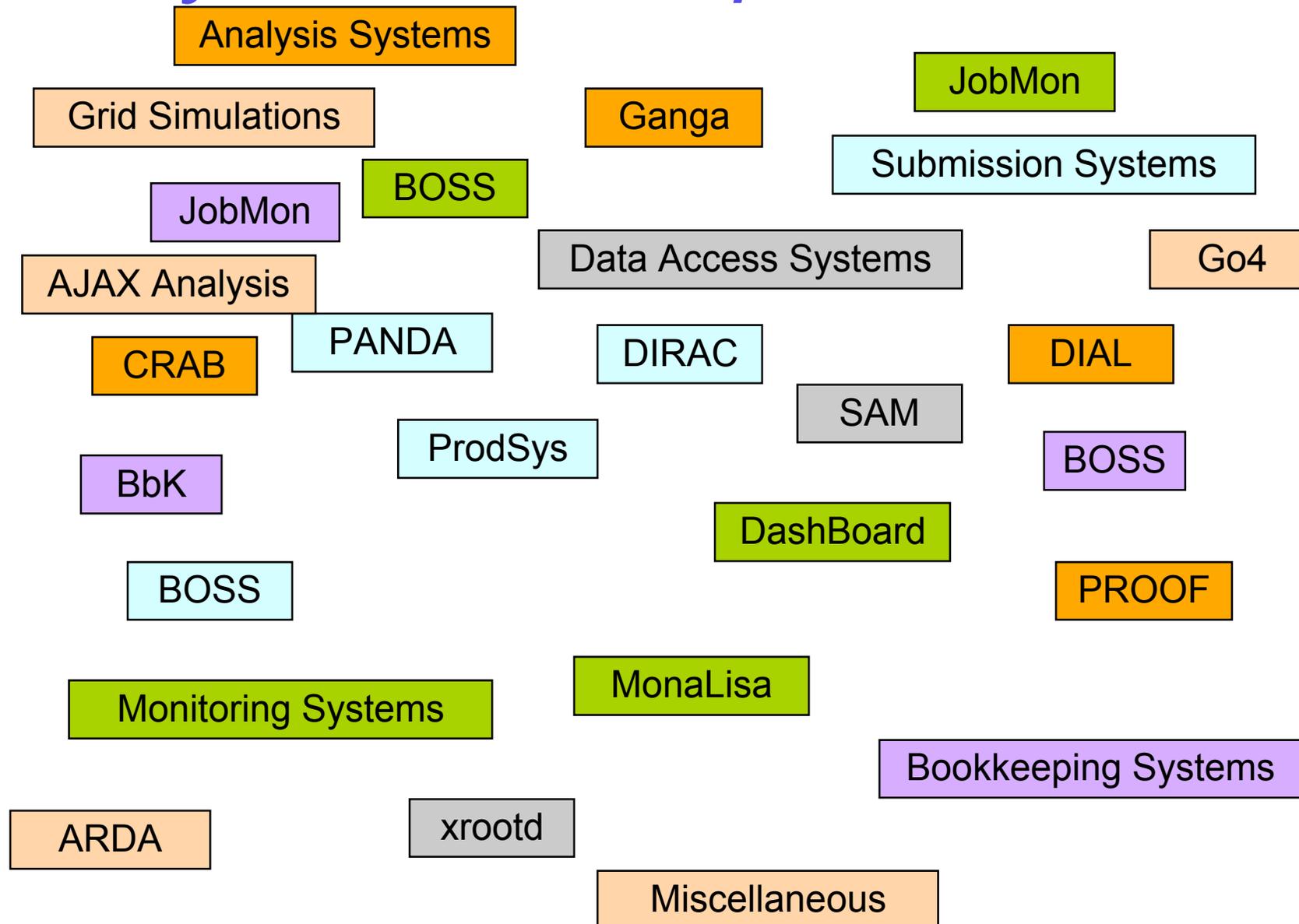
- Bilan succinct:
 - Bon résultats: Ça se consolide
 - Meilleure communication LCG/expériences (enfin :))
 - Mise en évidence des problématiques (stockage)



La production de données

- Bonne maîtrise de flots de travail pour la reconstruction et la simulation
- Bonnes performances
 - Passage à l'échelle
 - Robustesse
 - Meilleure compréhension
- Des tendances communes:
 - Chaque expérience a son système (voire plusieurs) système de production
 - Utilisation d'agents (LHCb, Atlas, CMS, Alice)
 - Stratégies opportunistes - « jobs pilotes »
 - Évidemment duplication d'efforts... :(

Analyse: Beaucoup d'efforts...



Tentative de classification

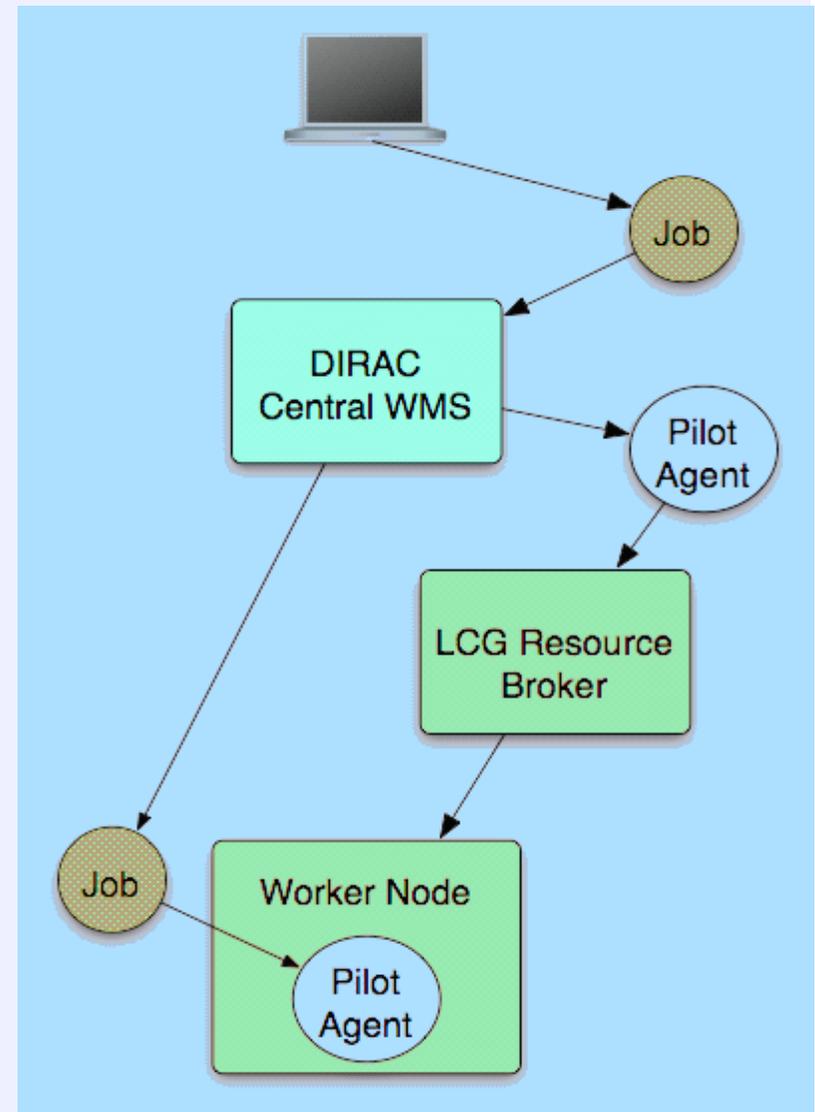
- Système d'analyse clients:

ALICE	ATLAS	CMS	LHCb
PROOF	DIAL GANGA	CRAB PROOF	GANGA

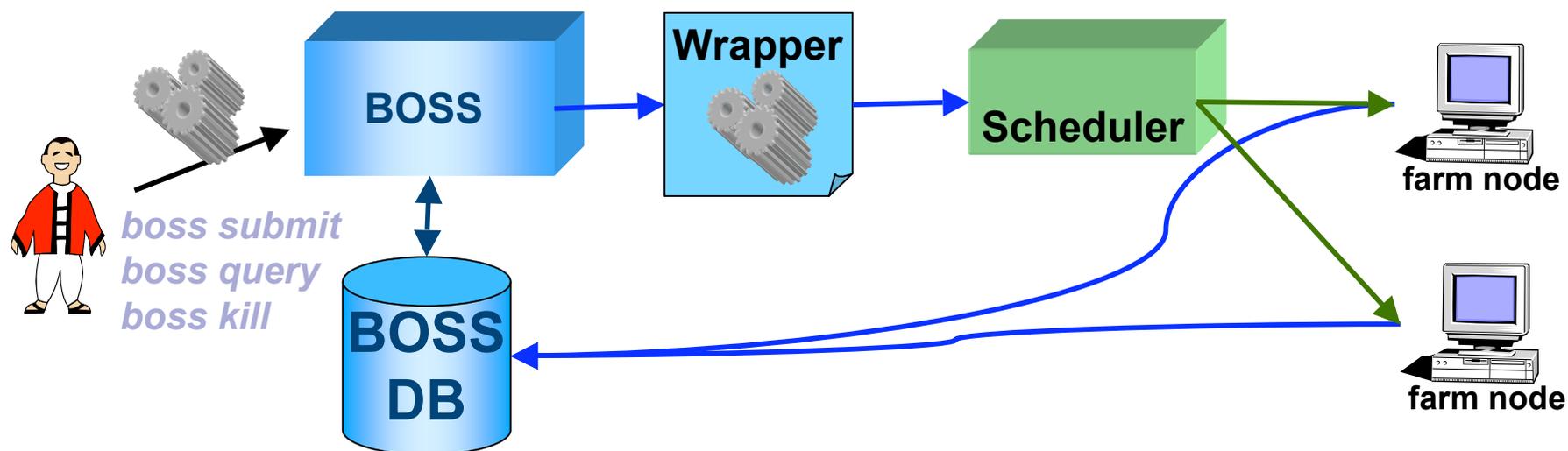
- Système de soumission (souvent les syst. de production):

ALICE	ATLAS	CMS	LHCb
AliEn (non présent)	ProdSys PanDA	BOSS	DIRAC

- The DIRAC Workload & Data Management System (WMS) is made up of **Central Services** and **Distributed Agents**
 - Realizes PULL scheduling paradigm
 - Agents are requesting jobs whenever the corresponding resource is available
 - Execution environment is checked before job is delivered to WN
 - Service Oriented Architecture masks underlying complexity



BOSS Workflow



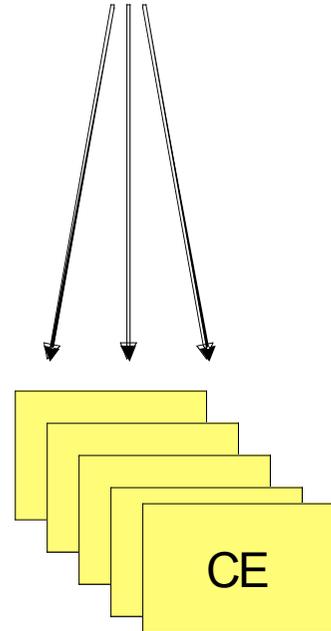
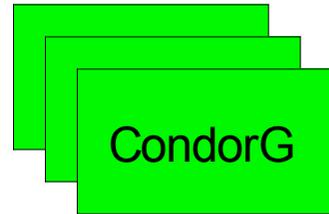
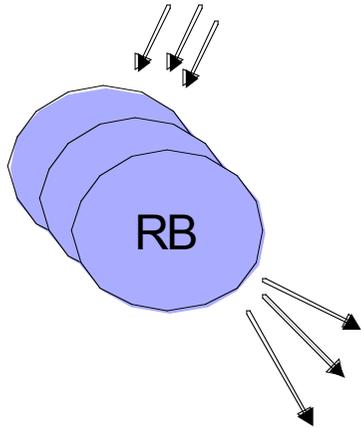
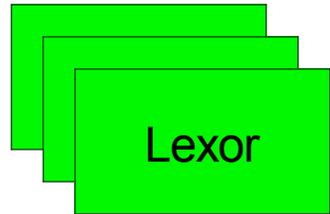
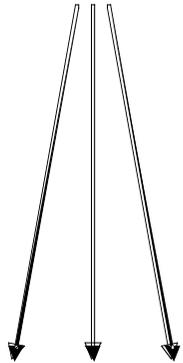
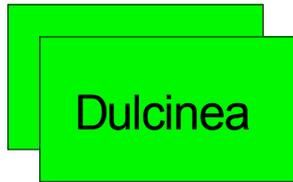
- User specifies job - parameters including:
 - Executable name.
 - Executable type - turn on customized monitoring.
 - Output files to retrieve (for sites without shared file system and grid).
- User tells Boss to submit jobs specifying scheduler i.e. PBS, LSF, SGE, Condor, LCG, GLite etc..
- Job consists of job wrapper, Real time monitoring service and users executable.

From Evolution of BOSS by Wakefield [240]



ATLAS Prodsys

ProdDB



What is Panda



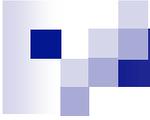
- ❑ PanDA – Production and Distributed Analysis system
- ❑ Project started Aug 17, 2005
- ❑ Baby Panda emerging!
- ❑ New system developed by U.S. ATLAS team
 - ❑ Rapid development from scratch
 - ❑ Leverages DC2/Rome experience
 - ❑ Inspired by Dirac & other systems
 - ❑ Already in use for CSC pre-production in the U.S.
 - ❑ Better scalability/usability compared to DC2 system
 - ❑ Will be available for distributed analysis users in few months
- ❑ One-stop shopping for all ATLAS users in the U.S.



Panda Core Components



- ❑ **Job Interface** – allows injection of jobs into the system
- ❑ **Executor Interface** – translation layer for ATLAS prodsys/prodDB
- ❑ **Task Buffer** – keeps track of all active jobs (job state is kept in MySQL)
- ❑ **Brokerage** – initiates subscriptions for a block of input files required by jobs (preferentially choose sites where data is already available)
- ❑ **Dispatcher** – sends actual job payload to a site, on demand, if all conditions (input data, space and other requirements are met)
- ❑ **Data Service** – interface to DQ2 Data Management system
- ❑ **Job Scheduler** – send pilot jobs to remote sites
- ❑ **Pilot Jobs** – lightweight execution environment to prepare CE, request actual payload, execute payload, and clean up
- ❑ **Logging and Monitoring systems** – http and web based
- ❑ **All communications through REST style HTTPS services** (via mod_python and Apache servers)



Conclusions

- La production de données est un exercice considérée comme maîtrisé
- Tout l'attention se focalise maintenant sur l'analyse des données
 - Extension des systèmes de production
 - Stratégie opportunistes/système surcouché
- Remarques personnelles:
 - Duplicité d'efforts: la roue réinventée plus d'une fois
 - Les expériences communiquent avec LCG mais pas entre elles - trop tard pour factoriser les efforts/standarts ?
 - Nous ne sommes pas encore prêt pour l'analyse
 - Beaucoup de points noirs: l'équité entre utilisateurs, temps de réponse minimal, gestion des données
 - 2007 c'est demain !!! Et ils sont où les physiciens ?



Références

- <http://www.tifr.res.in/~chep06/>
- <http://indico.cern.ch/conferenceTimeTable.py?confId=048>
- <http://indico.cern.ch/sessionDisplay.py?sessionId=8&slotId=0&confId=048#2006-02-13>
- <http://indico.cern.ch/sessionDisplay.py?sessionId=9&slotId=0&confId=048#2006-02-13>