

➤ Activités du computing LHCb depuis le redémarrage du LHC

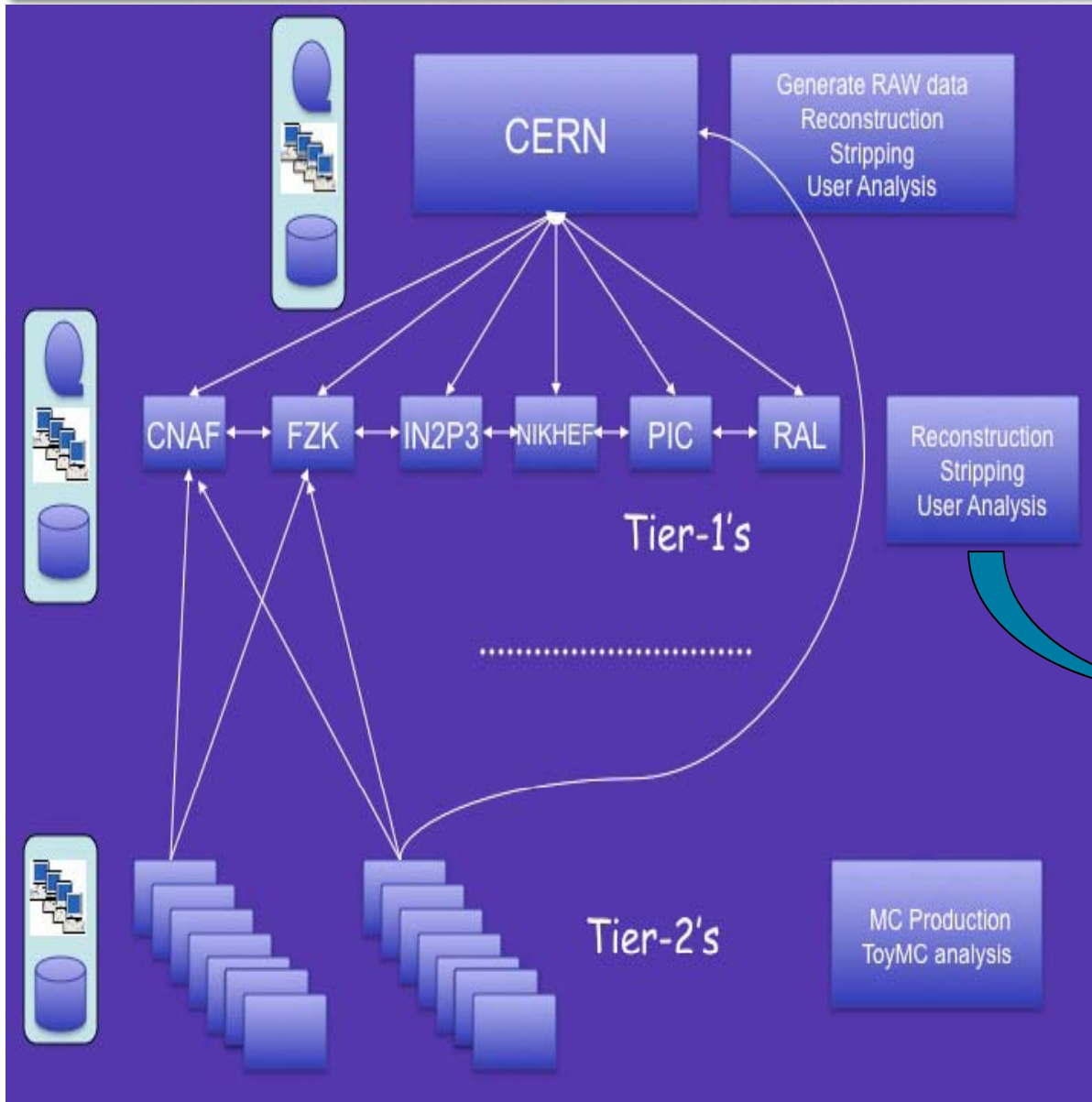
Activités de LHCb au T1 – IN2P3

Luisa Arrabito

LCG France au CPPM le 24/25 juin 2010



Rappel du Computing Model



- **30 mars 2010 :**
premières collisions
- **31 mars 2010 :**
premières données disponibles pour l'analyse
- **Toute la chaine appliquée aux collisions 2010**



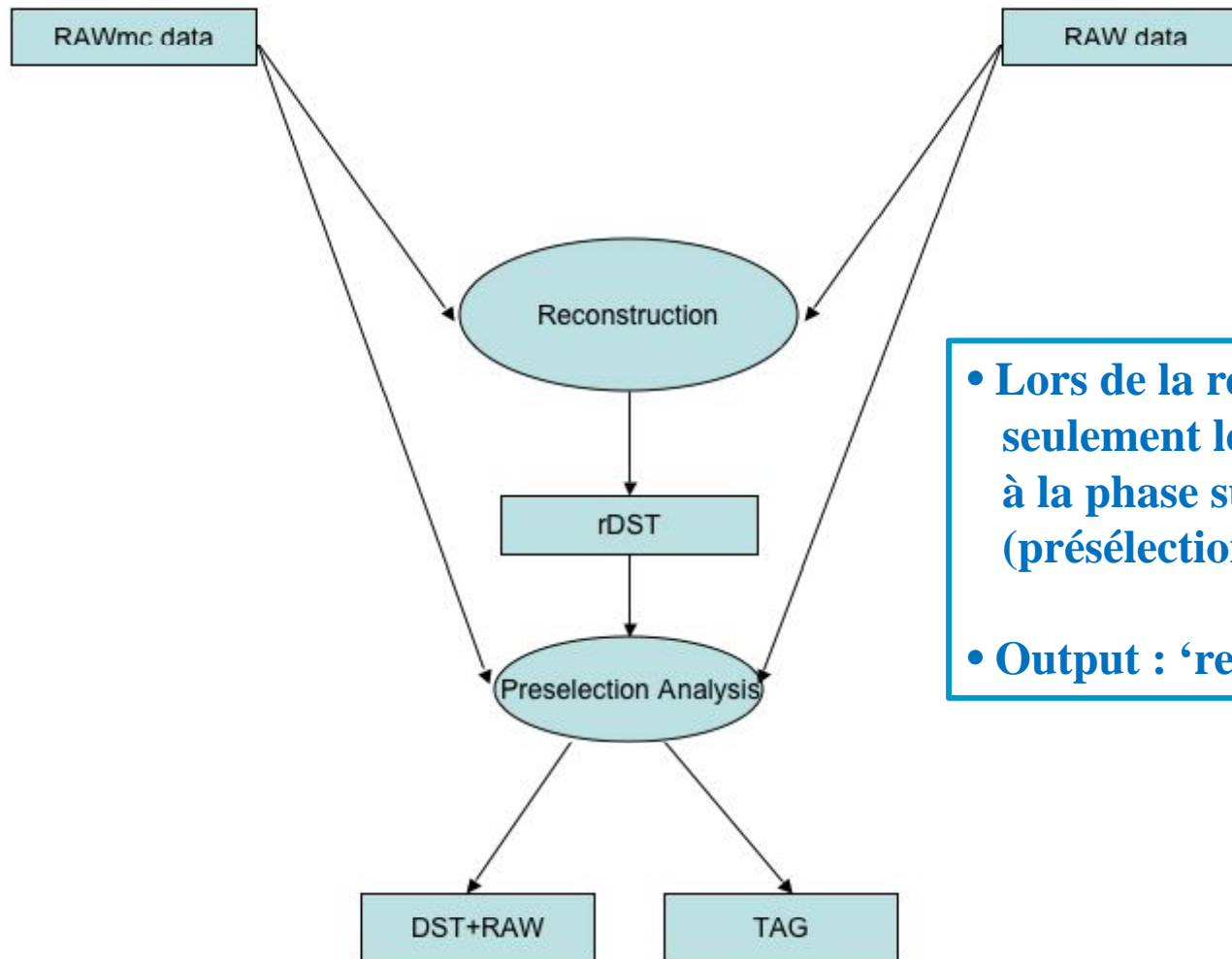
➤ **Activités liées aux collisions 2010**

- **Reconstruction**
- **Stripping**
- **Analyse**
- **Validation de nouveaux workflows**

➤ **Autres activités**

- **Stripping des données des collisions 2009**
- **Simulation MC**
- **Analyse**

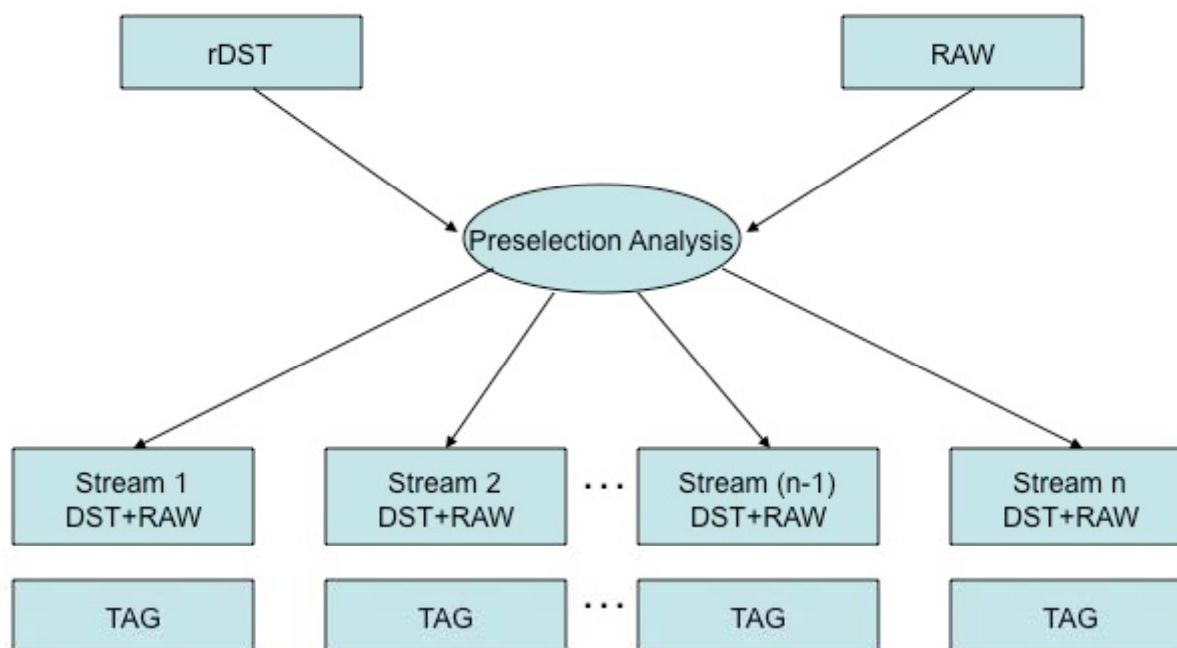
Reconstruction des RAW data -> T1s



- Lors de la reconstruction on garde seulement les informations nécessaires à la phase successive de 'stripping' (présélection)
- Output : 'reduced' DST (rDST)

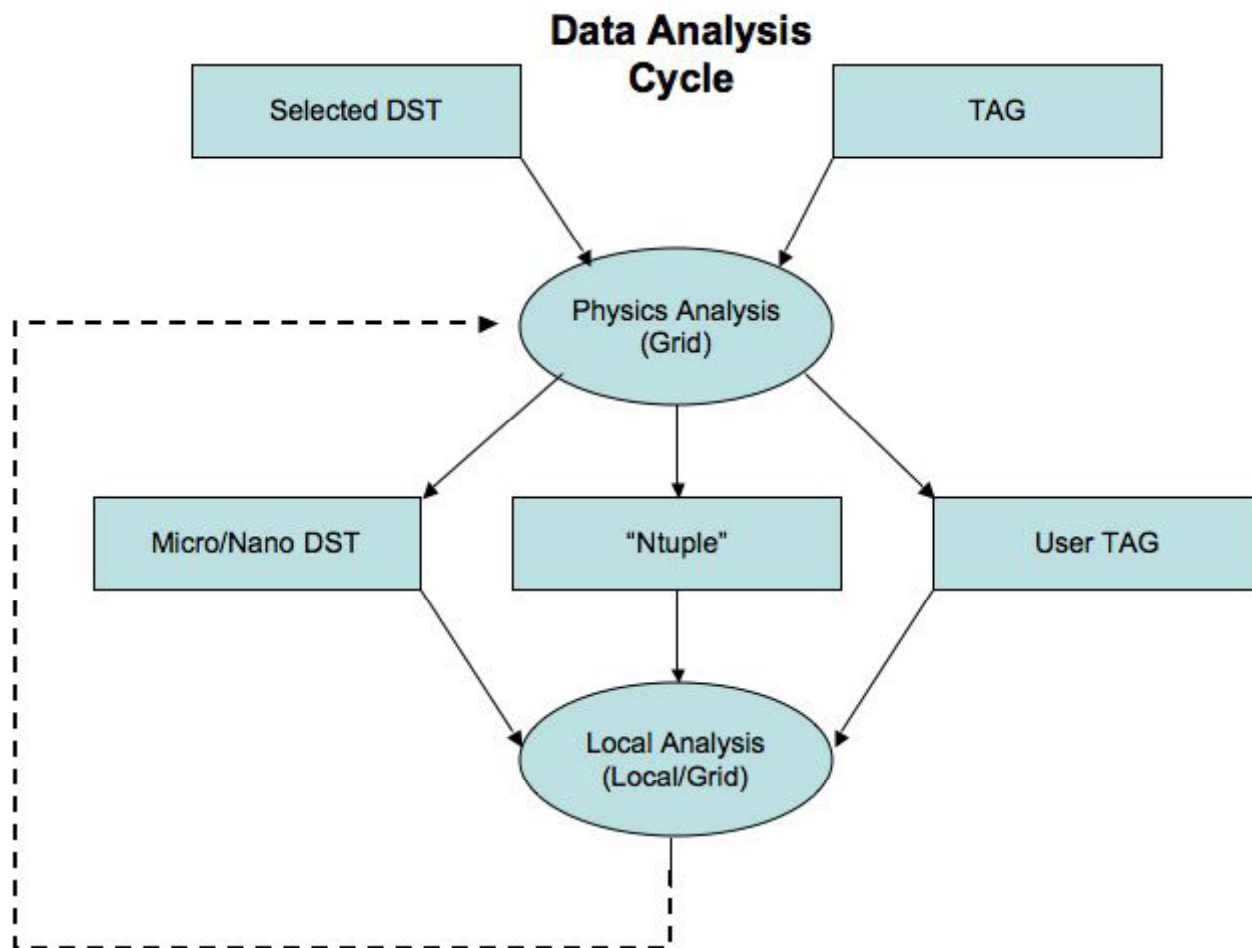


Stripping (présélection) -> T1s



- Les événements sont présélectionnés selon “n” streams pour l’analyse finale
- Output : (full) DST (contenant des infos de reconstruction supplémentaires) + RAW data associées

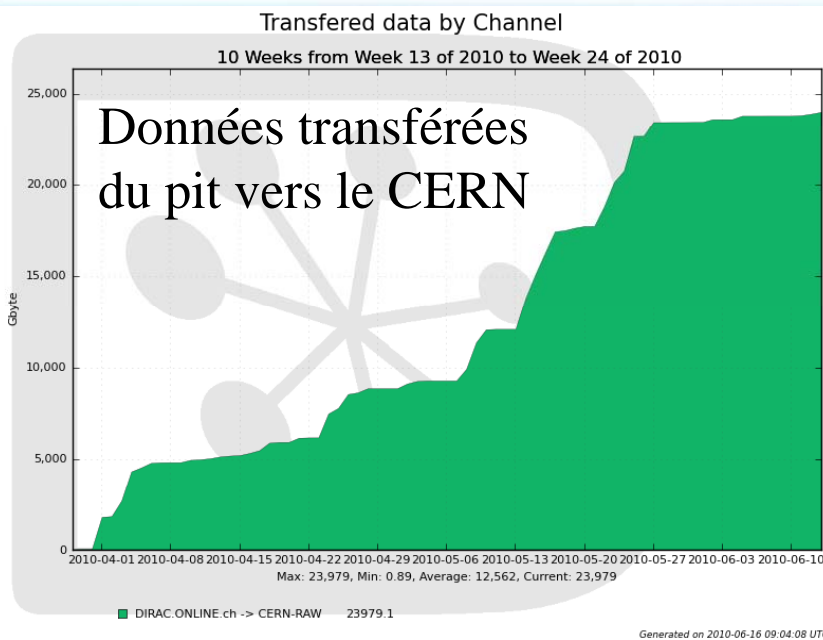
Analyse -> T1s



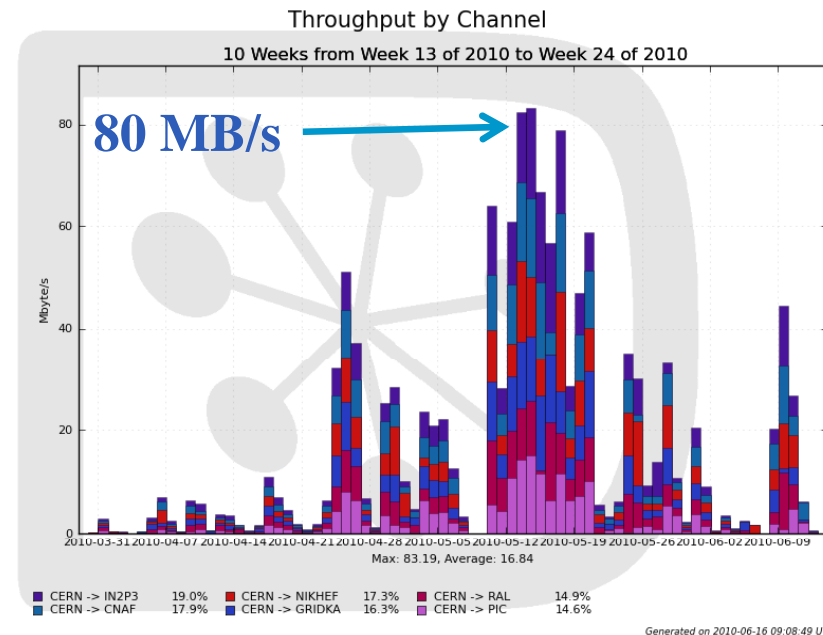
- L'analyse se fait essentiellement à partir de l'output du stripping
- Output : Ntuple/MicroDST



Transfères depuis le 30 mars 2010



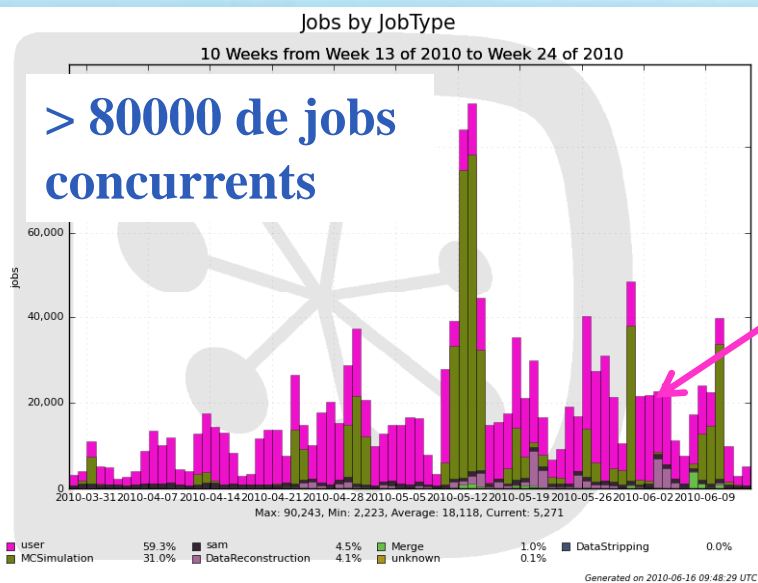
Transfères du CERN vers les T1s



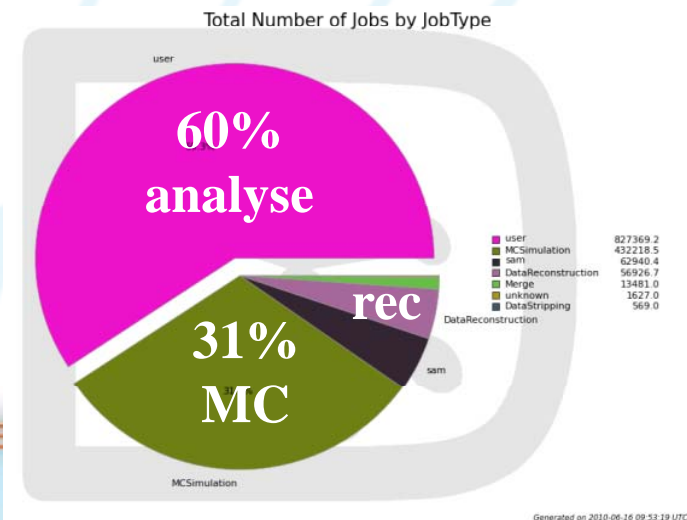
Distribution des RAW data dans les T1s

DIRAC SE	Size (TB)	Files
CERN-RAW	30.5	20688
CNAF-RAW	2.1	1710
GRIDKA-RAW	2.9	2185
IN2P3-RAW	3.4	2736
NIKHEF-RAW	4.0	2933
PIC-RAW	0.9	803
RAL-RAW	3.2	2387

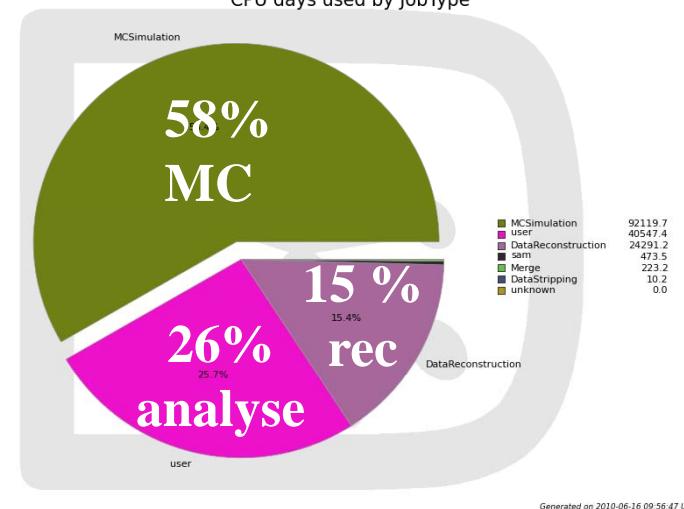
Jobs et CPU depuis le 30 mars 2010



- Analyse
- Simulation MC
- Reconstruction
- SAM
- Merge
- Stripping



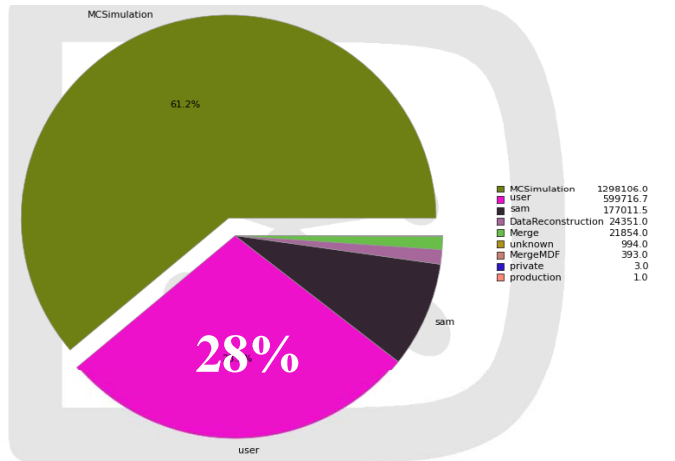
CPU par type de job





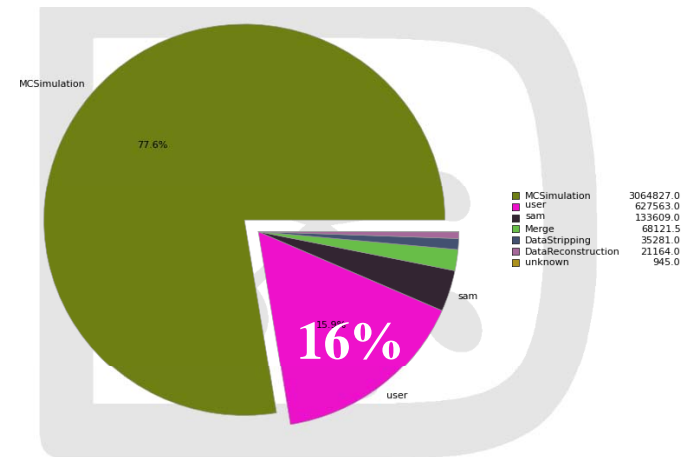
Historique des jobs d'analyse depuis 2009

Premier semestre 2009



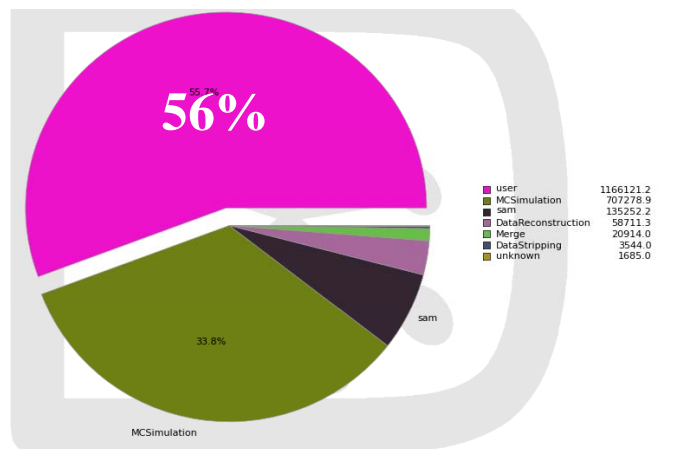
Generated on 2010-06-18 15:33:35 UTC

Deuxième semestre 2009



Generated on 2010-06-18 15:38:36 UTC

Premier semestre 2010



Generated on 2010-06-18 15:40:50 UTC

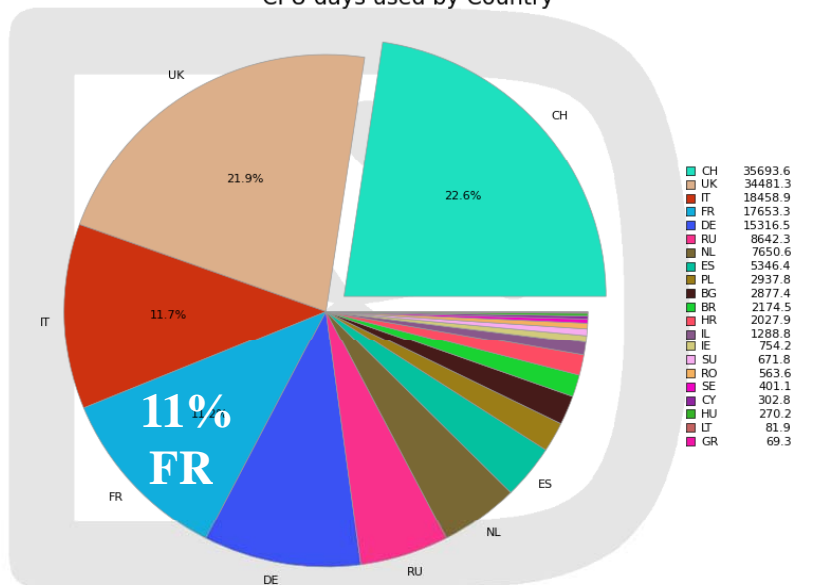
➤ **Total des jobs :**
2.1 M / 3.9 M / 2.1M

➤ **Jobs d'analyse :**
600K / 628K / 1.2M

Jobs et CPU depuis le 30 mars 2010

CPU utilisé par pays

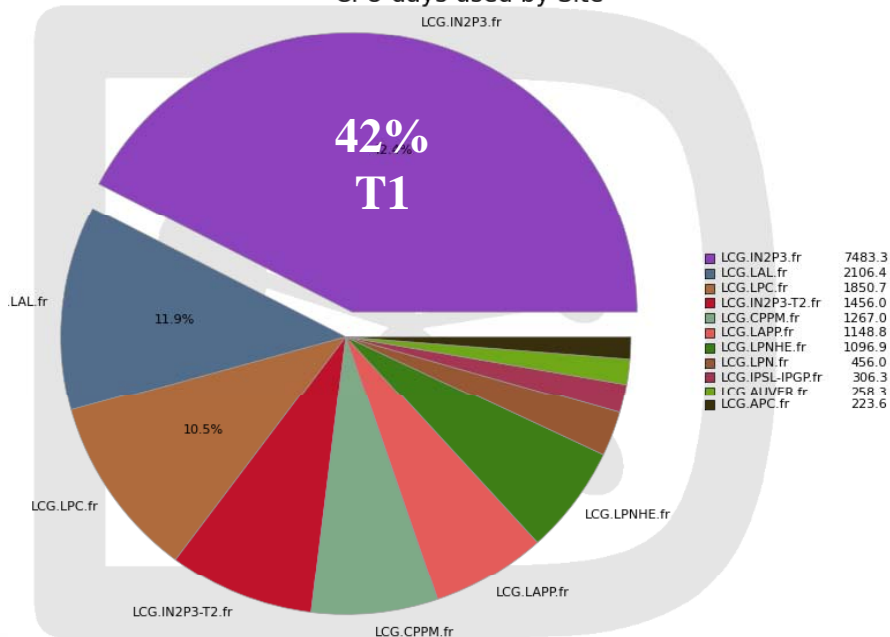
CPU days used by Country



Generated on 2010-06-15 10:13:36 UTC

En France

CPU days used by Site

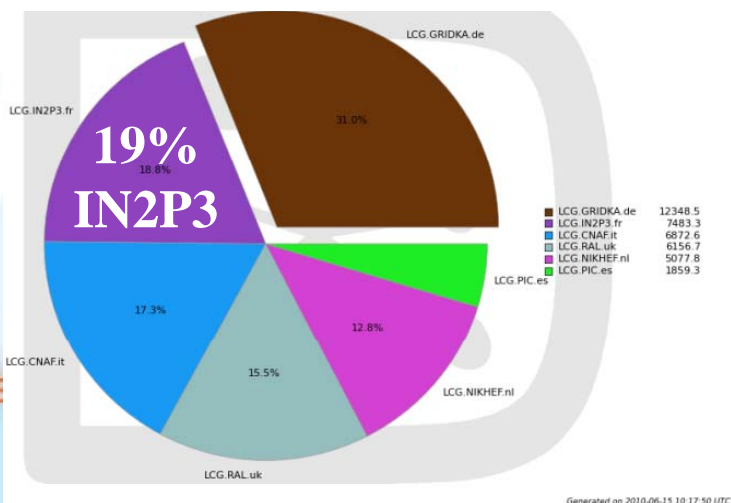


Generated on 2010-06-15 10:23:40 UTC

Engagements 2010 des T1s pour LHCb

Site	CPU (HEPSPEC06)	Disk (TB)	Tape (TB)
CERN	23000 (100%)	1290 (100%)	1800 (100%)
CNAF	5500 (13%)	450 (14%)	450 (19%)
FZK	7480 (17%)	560 (17%)	408 (17%)
IN2P3	9742 (22%)	728 (22%)	531 (22%)
NL-T1	8992 (20%)	707 (21%)	1012 (42%)
PIC	2632 (6%)	197 (6%)	189 (8%)
RAL	8184 (19%)	612 (19%)	446 (19%)
TOT	(97%)	(99%)	(127%)

CPU utilisé dans les T1s



- Shares non respectées
- Comment rendre notre T1 plus attractif?
- **Remarques :**
 - Utilisation encore faible par rapport aux demandes
 - 'Classement' fluctue dans le temps

Problèmes majeurs aux T1s

➤ Accès aux fichiers

- L'utilisation de xroot prise en compte -> Utilisé actuellement au CERN
- Solution actuelle pour les jobs de production : download des fichiers
- Tuning des paramètres dans les sites dCache

➤ CPU/wallclock

- Inefficacités observées par les sites
- Trop de pilotes soumis ('filling mode') (le pilot commet un suicide si aucune tâche n'est disponible, mais ceci après quelques minutes)
- Connexions 'en suspens' (data upload, data-stream avec les serveurs dcap/root, storage shortage, AFS outages, jobs hanging)

➤ Shared Area

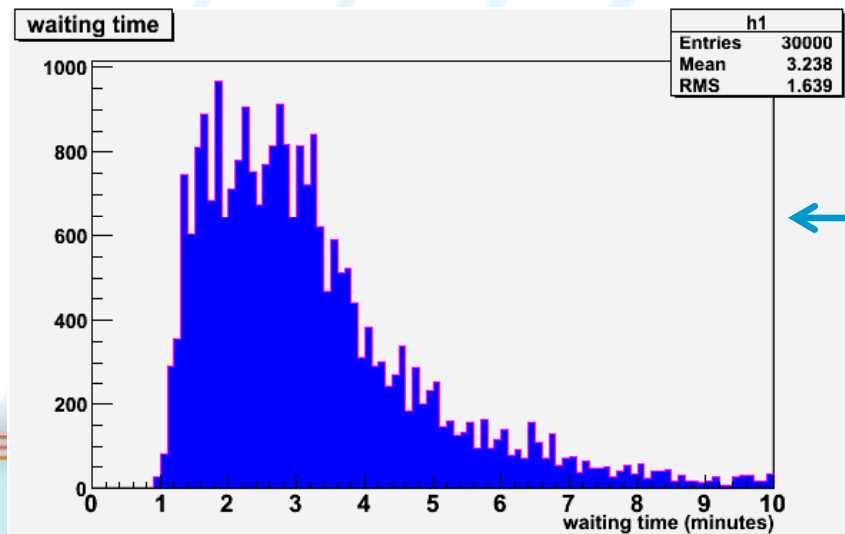
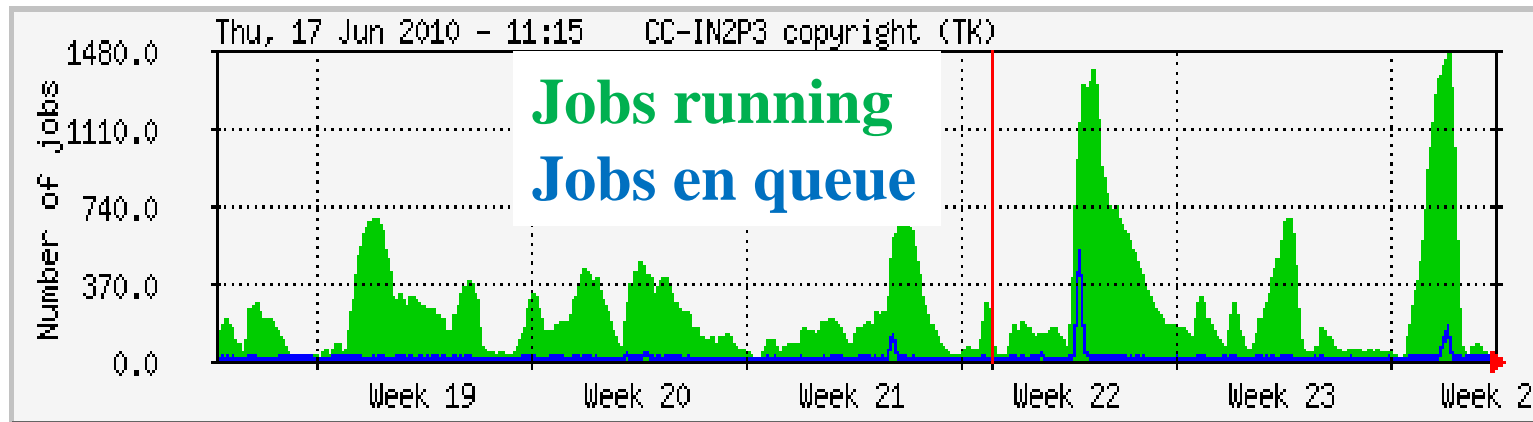
- Service critique pour lhcb (zone contenant le software de l'expérience)
- Ex. Timeouts pendant le source de l'environnement du job



Activité LHCb au T1-IN2P3

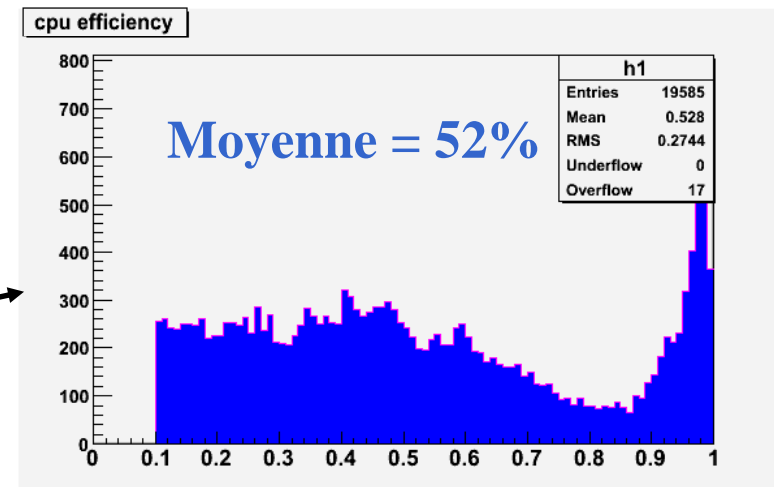
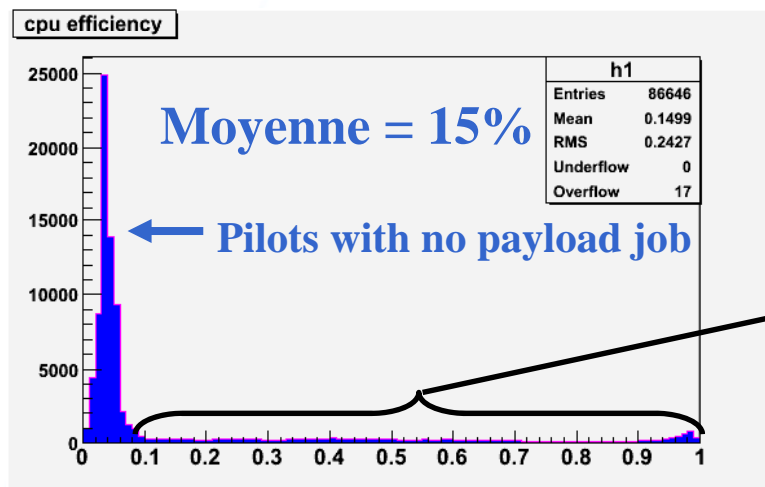


Mai-Juin 2010



**Temps d'attente dans la queue du système de batch :
~ 3 minutes**

Effacité CPU/WallclockTime Janvier – Mars 2010

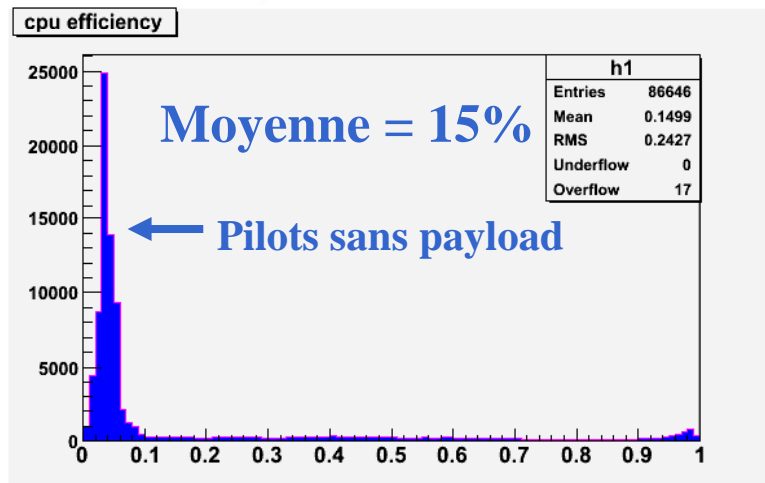


- Très faible efficacité CPU pour les jobs lhcb -> 15%
- 77% des jobs a une efficacité cpu < 10% -> Pilotes sans payload
- Mauvaise utilisation des ressources -> 506 h (21 jours) de CPU pour des pilotes sans payload job en 3 mois

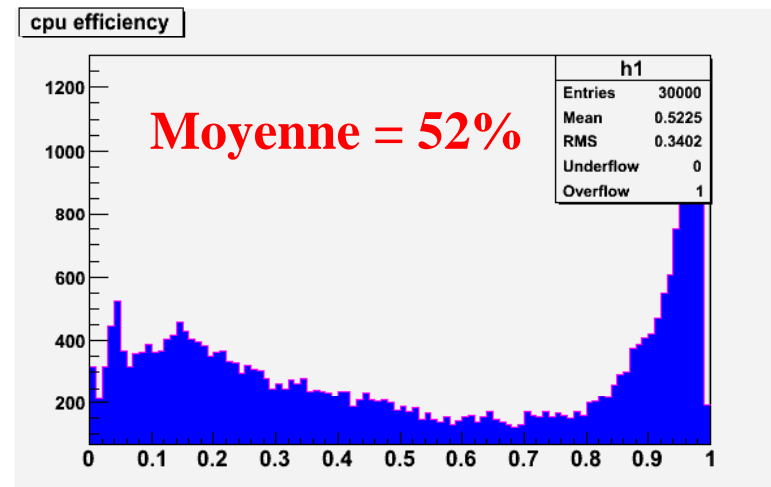


Effacité CPU/WallclockTime

Janvier – Mars 2010



Mai – Juin 2010



- Optimisation de la soumission des pilotes
- Nette amélioration de l'efficacité CPU
- Une ultérieure amélioration est possible (accès aux fichiers, connexions qui restent en suspens etc...)

▶ Evolutions récentes au T1-IN2P3 pour LHCb



- **2 CEs CREAM mis en production (1 pour le T1 et 1 pour le T2)**
- **Tests sur glexec effectués en collaboration avec LHCb**
- **Ouverture des portes dCap à LHCb (problèmes récurrents avec gsidcap)**
- **Protocol xrootd prêt à être utilisé**
- **Nouvelle VO box en SL5 mise en production**
- **Publication du facteur de normalisation du CPU, CPUScalingReferenceSI00**
- **Reconfiguration et ajout de nouvelles queues (en particulier pour les jobs de reconstruction)**
- **Système de protection pour les lectures bande testé et prochainement mis en production**

Incidents majeurs au T1-IN2P3

➤ Shared area

- Timeout pendant le source de l'environnement des jobs (taux d'erreur : 1- 10%)
- Plusieurs causes possibles (charge serveur AFS, client AFS, comportement anormal des wns en cas de forte I/O -> gèle de machines)
 - > Résolu pour les jobs utilisateurs après changement de type et version du serveur AFS
 - > Problème pas complètement cerné (SAM tests en échec de temps en temps)

➤ Autres

- Instabilités SRM détectées par les SAM jobs
- Signaux envoyés par BQS aux jobs qui atteignent la limite d'une queue non gérés par les applications de lhcb
- Incident AFS le 26-27 avril
- CE CREAM : le status des jobs n'est pas reporté correctement par le WMS jobs sont vus 'scheduled' par le WMS alors qu'ils ne sont plus dans le système
 - > En attente d'un upgrade du WMS qui corrige le problème
- Erreur Oracle -> résolu (augmentation du max mémoire dans la Shared Global Area)



Conclusions

- **LHCb a validé toute la chaîne du computing model avec les données des collisions 2010 : de la réception des RAW DATA, jusqu'à l'analyse faite par les physiciens**
- **Rôle important des T1s dans cette phase -> bilan plutôt positif pour le T1 français**
- **Plusieurs canaux de communication entre la VO et les T1s (en plus du meeting journalier pour les opérations de LHCb)**
- **Capacité de LHCb et du T1-IN2P3 à s'adapter rapidement en cas d'incidents ou d'imprévus**