

# Support pour LHCb au CC-IN2P3

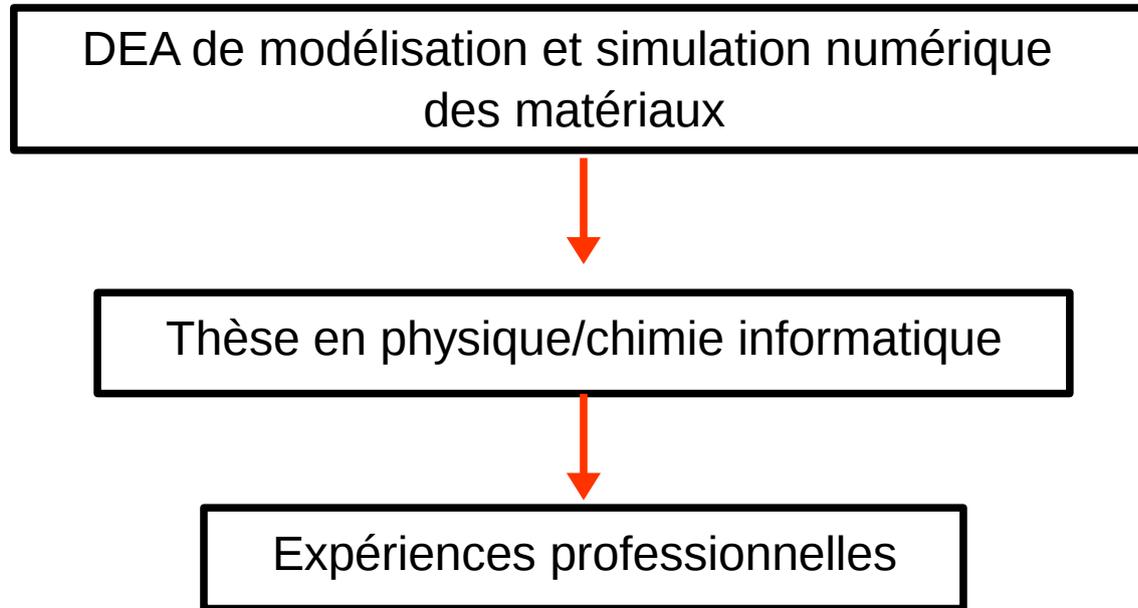
Adil Khalifa, Journée LCG-France, 19 Juin, Villeurbanne

DEA de modélisation et simulation numérique  
des matériaux

```
graph TD; A[DEA de modélisation et simulation numérique des matériaux] --> B[Thèse en physique/chimie informatique]; B --> C[Expériences professionnelles];
```

Thèse en physique/chimie informatique

Expériences professionnelles

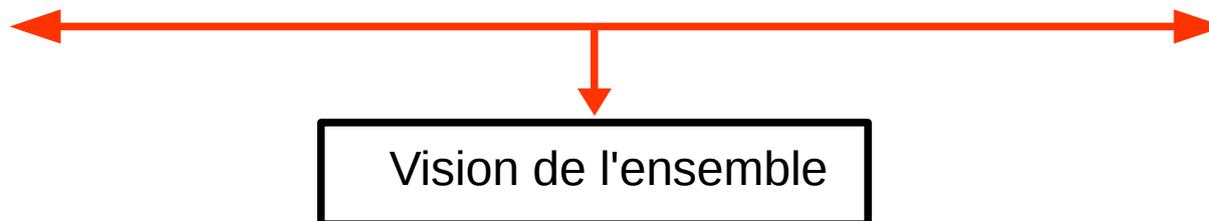


## Support en calcul scientifique

- IR2 en calcul scientifique au CNRS/CCIN2P3
- Ingénieur en informatique scientifique au CINES

## Recherche en HPC

- Stage au CEA/Saclay
- Thèse à l'université lorraine
- ATER à l'université lorraine
- IR2 à l'ENSIC





- T1 : 7 sites

UK : RAL

Italy : CNAF

Germany : GRIDKA

France : IN2P3

Netherlands : NL

Russia : RRCKI

Spain : PIC

Sites/2018	CPU (HS06)*k	Disk (PB)	Tape (PB)
RAL	74,82	7,32	18,231
CNAF	46,805	5,61	11,4
GRIDKA	42,25	4,1	10,27
IN2P3	30,4	3,2	7,300
NL	24,6	2,31	3,993
RRCKI	16,4	2,3	3
PIC	14,8	1,43	2,668

**Source : REBUS**

# Les pledges de T0

- T1 : 7 sites

UK : RAL

Italy : CNAF

Germany : GRIDKA

France : IN2P3

Netherlands : NL

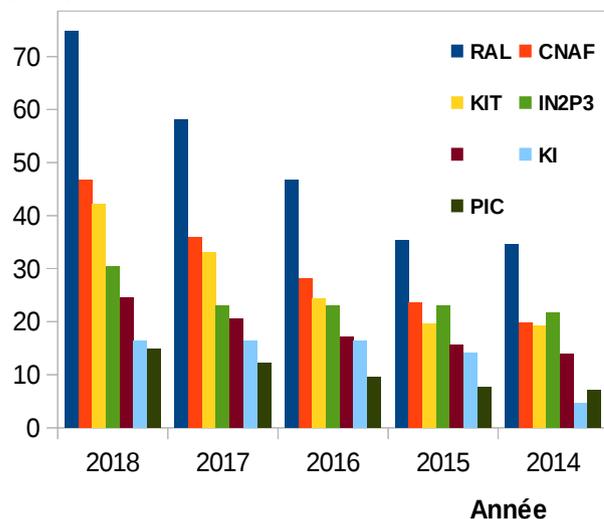
Russia : RRCKI

Spain : PIC

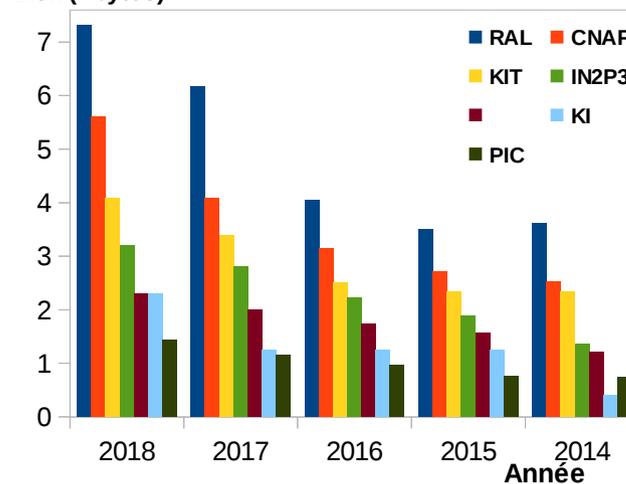
Sites/2018	CPU (HS06)*k	Disk (PB)	Tape (PB)
RAL	74,82	7,32	18,231
CNAF	46,805	5,61	11,4
GRIDKA	42,25	4,1	10,27
IN2P3	30,4	3,2	7,300
NL	24,6	2,31	3,993
RRCKI	16,4	2,3	3
PIC	14,8	1,43	2,668

Source : REBUS

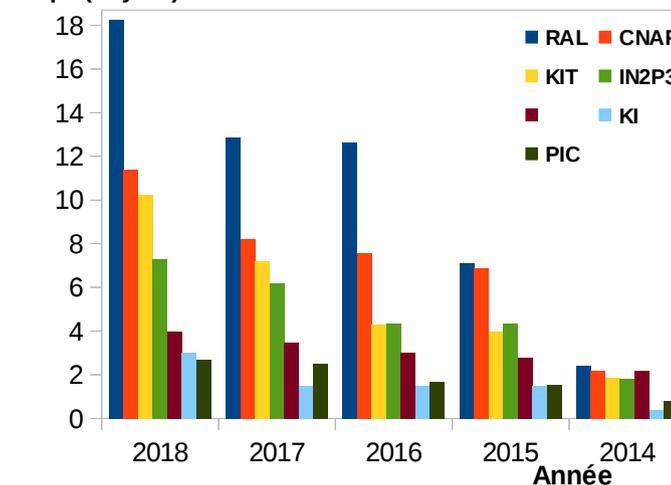
CPU (hs06)\*k



Disk (Pbytes)

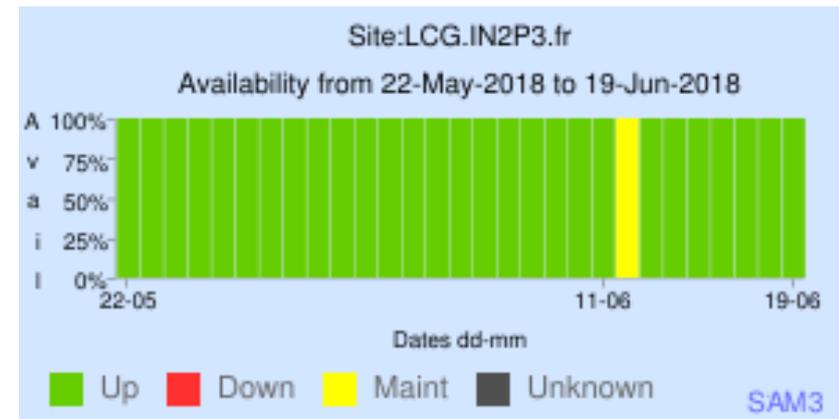
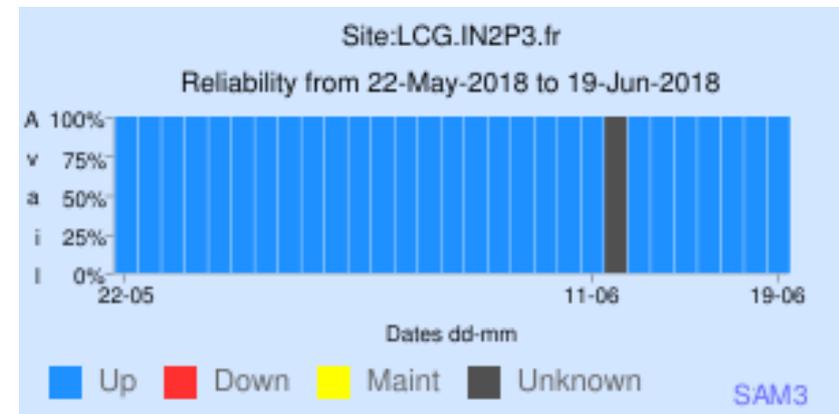


Tape (Pbytes)

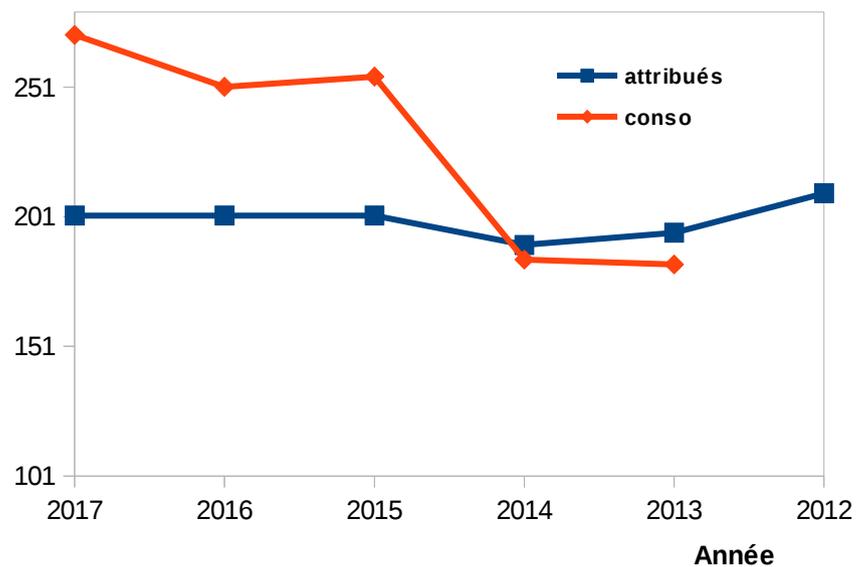


- **Deux CREAMs** :  
cccreamceli02  
cccreamceli06
- **Batch** : UGE
- **Deux queues de batch** :  
Longlasting : 168h ; 16G  
Long : 48h ; 16G → actif
- **3 projets** :  
lhcb, lhcb\_prod, lhcb\_pilot -> actif
- **1 VOBOX**
- Migration vers CL7 depuis 02/18.
- **Réseau** :  
LHC-OPEN 20G/s : entre CC et CERN : RAW  
LHC-ONE 30G/s : pour d'autres transferts

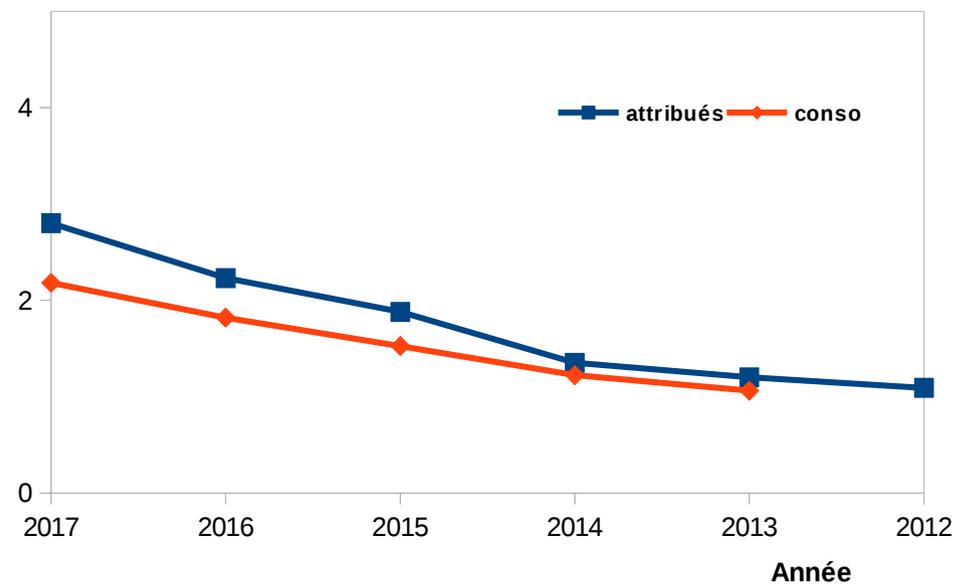
Space	Used	Free	Size	N. files
LHCb_Disk	2.1 PB	884TB	3.0PB	1.110.523
LHCb_Tape	42GB	26TB	27TB	1.780.851
LHCb_User	104TB	27TB	132TB	1.731.181



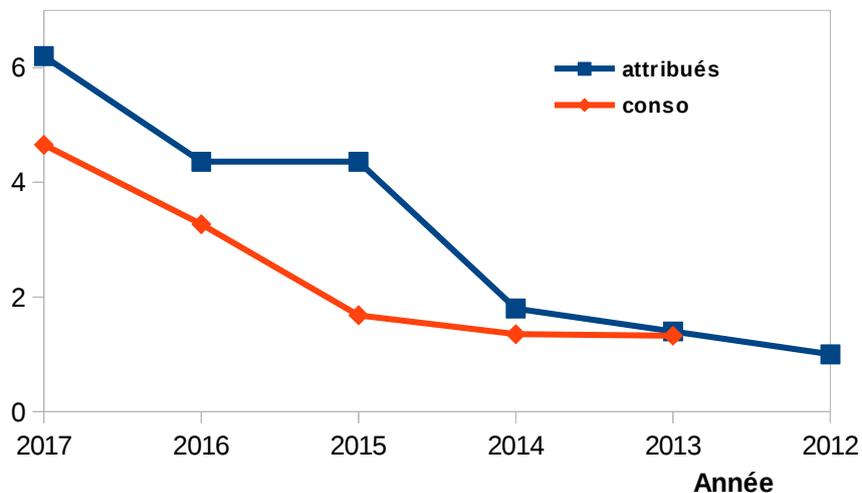
CPU (hs06)\*M



Disk (PB)



Tape (PB)

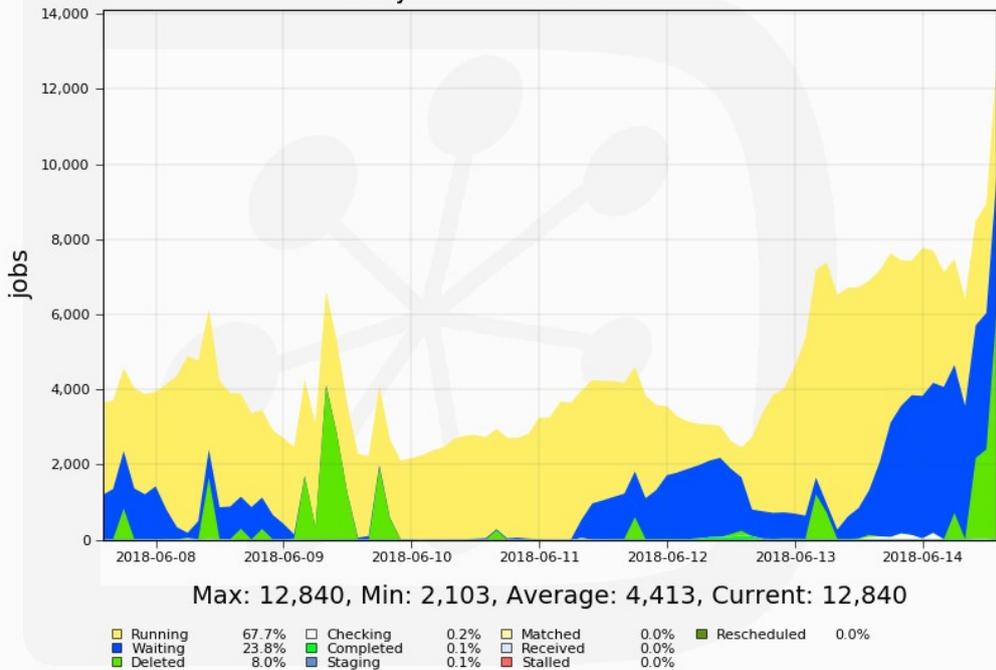


Bon rapport entre les consommations et les pledges

- **Identique par rapport aux autres expériences :**
  - Coordonner l'activité de LHCb (calcul, transfert, stockage)
  - S'interfacer entre LHCb et les équipes informatiques CCIN2P3
  - Utiliser les différents canaux de communication
  - mettre en place les nouveaux projets de l'expérience (MFJ)
- **Au parallèle :**
  - Participer à l'activité opérationnelle de LHCb au CERN (Shift)
  - Participer à l'enseignement à l'UCBL
  - Assurer une veille technologie

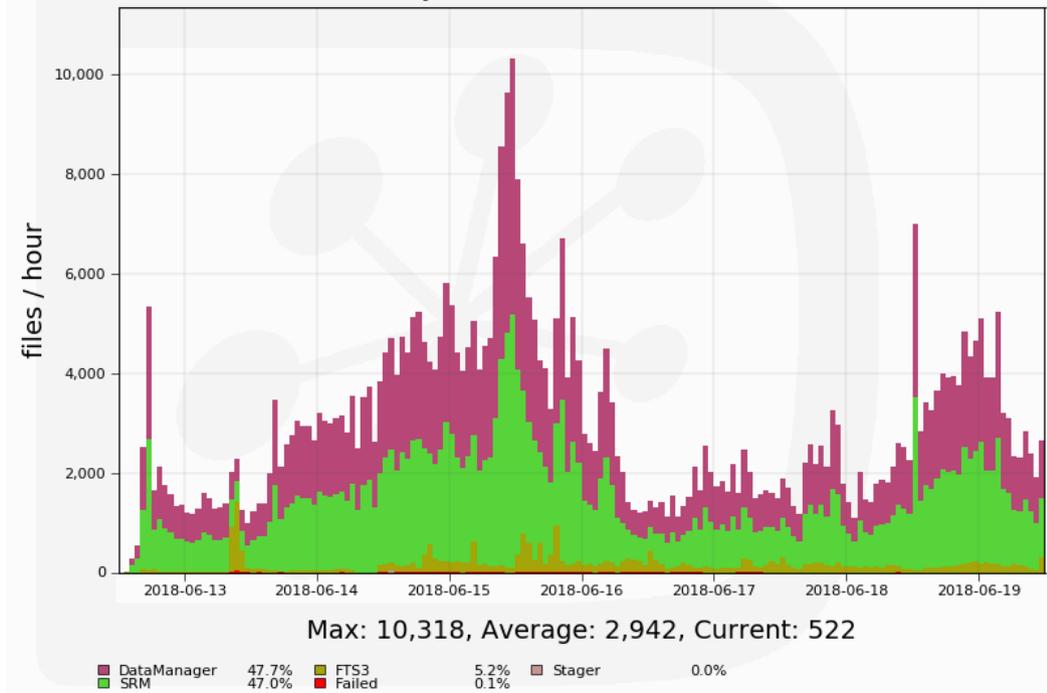
- **Dirac** : monitoring, accounting de productions, de jobs, de données, de ressources informatiques.

Jobs by Status  
7 Days from 2018-06-07 to 2018-06-14



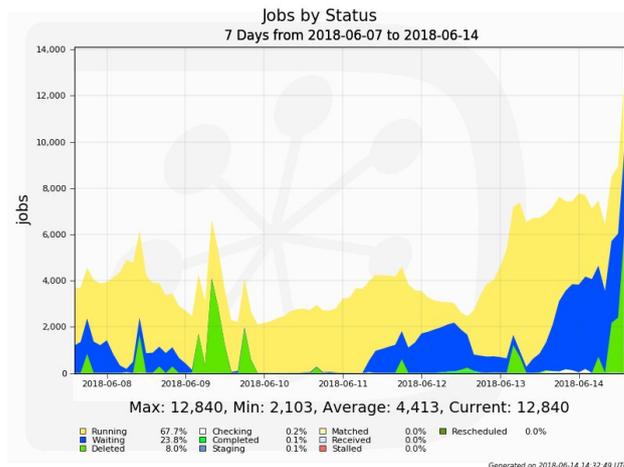
Generated on 2018-06-14 14:32:49 UTC

successful Transfers by Protocol (CCIN2P3 as Destination SE)  
7 Days from 2018-06-12 to 2018-06-19

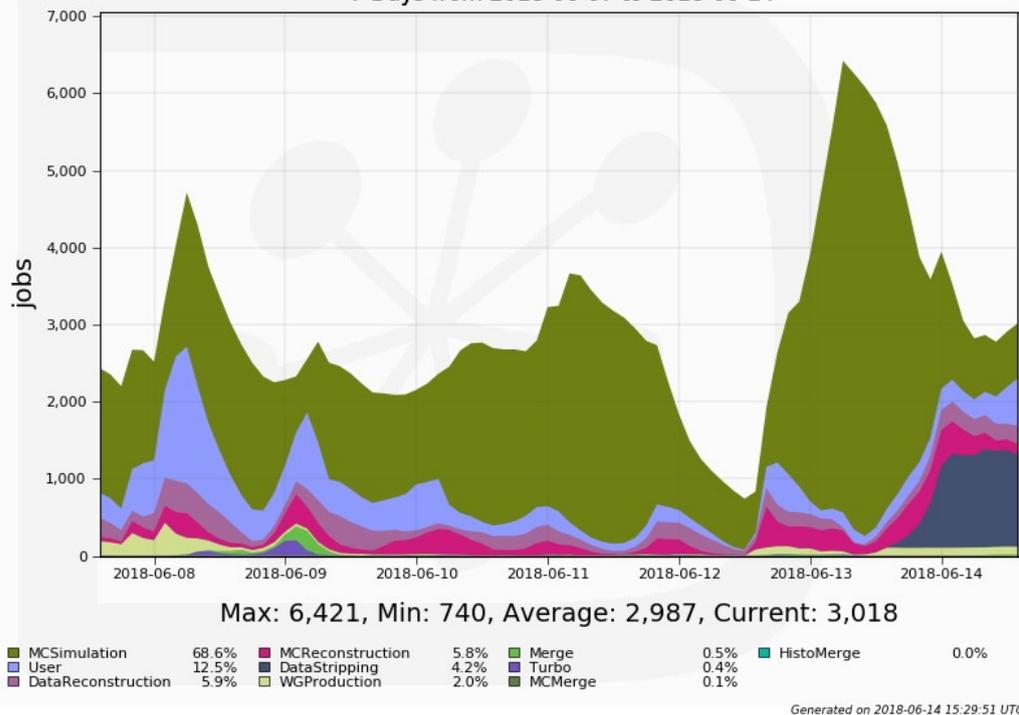


Generated on 2018-06-19 12:12:32 UTC

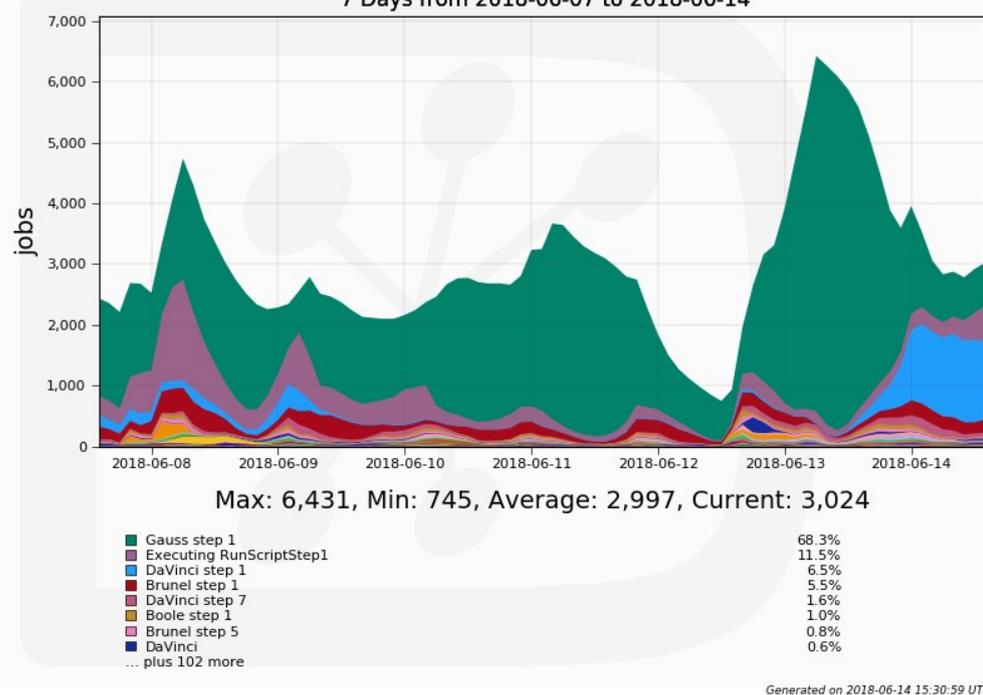
- Dirac : monitoring, accounting,



Running job by JobSplitType  
7 Days from 2018-06-07 to 2018-06-14



Running job by MinorStatus  
7 Days from 2018-06-07 to 2018-06-14



- **D'autres outils pour monitoring ?**

**outils internes**, développés au ccin2p3: mrtguser et scripts.

test **SAM LHCb** : disponibilité, faillibilité

**CheK\_MK** : machines et services

**Dashboard lhcb** : stockage en ligne

**fts3-CERN** : transferts

**GOCDDB** : downtime

**Grafana-CERN** : réseau

- **Canaux de communications ?**

Réunions hebdomadaires, Elog, GGUS, email, Shift

- **Y-a-t-il une coordination nationale des sites ?**

Non, Pas besoin, LCGFR est largement suffisant

Par contre, Oui pour renforcer la présence LHCb au sein LCGFR

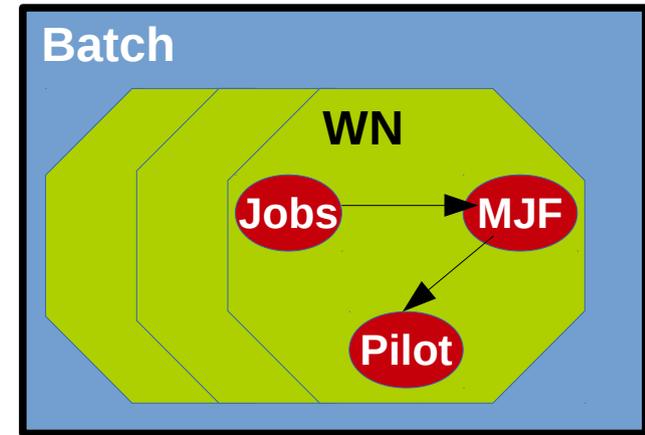
- **LHCb semble être très bon pour rentrer dans les fermes batch ?**

Oui, au ccin2P3, les objectifs lhcb sont remplis très vite que les autres VO

Parce que : les jobs sont mono-cœurs.

## Définition

MJF est un mécanisme informatique, qui sera exécuté une seule fois par un job afin de fournir au pilot-job des informations spécifiques (WN /job).



**But** : meilleure utilisation des ressources CC

## Stratégie

- Deux variables d'environnement :

**MACHINEFEATURES** : Pointe vers le répertoire contenant des fichiers spécifiques au WN.

total\_cpu, hs06, shutdowntime, grace\_secs

Ex : `$MACHINEFEATURES/total_cpu`

**JOBFEATURES** : Pointe vers le répertoire contenant des fichiers spécifiques aux jobs.

job\_id, allocated\_cpu, jobstart\_secs, hs06\_job, max rss\_bytes,....

Ex : `$MACHINEFEATURES/jobstart_secs`

**Implémentation** : HTCondor, UGE, Slurm, LSF, Torque/Pbs

Au CCIN2P3 : UGE

- **MACHINEFEATURES** :

hostname, total\_cpu, jobslots, phys\_core, factorscaling, hs06.

- **JOBFEATURES** :

job\_id, allocated\_cpu, jobstart\_secs, scratch\_limit\_bytes, cpufactor\_lrms, cpu\_limit\_secs, cpu\_limit\_secs\_lrms, priority\_factor, max\_rss\_bytes, max\_swap\_bytes, mem\_limit\_MB, wall\_limit\_secs, hs06\_job,

## Perspectives

Simulation rapide

Utilisation des multi processeurs CPU

Développer des codes parallèles

Chercher la performance de la machine

## Conséquences

Production de plus en plus de données → capacité et rapidité

Calcul haut débit avec performances → Machines HTC avec perf

## Questions

Avez-vous une idée quantitative sur le besoin en CPU et en stockage ?

Que pensez-vous pour les machines utilisées dans l'avenir :

HTC, ± HPC ou hybride (HTC+HPC=MTC)

**Pour les sites** : Sommes-nous prêts à répondre à ce besoin réel et inévitable ?

La grille de demain utilisera des probablement des machines HPC

La grille de demain utilisera des probablement des machines HPC

Un nouvel avancement en calcul scientifique

## Supercalculateur SUMMIT, USA



IBM avec Nvidia  
200 pétaflops  
4608 nœuds  
réseau InfiniBand à 200Gbit/s

**Chaque nœud :**  
2 CPU IBM Power9  
6 GPU Nvidia Tesla Volta  
2 T de memoire

La grille de demain utilisera des probablement des machines HPC

Un nouvel avancement en calcul scientifique

**Supercalculateur SUMMIT, USA**



Merci de votre attention

IBM avec Nvidia  
200 pétaflops  
4608 nœuds  
réseau InfiniBand à 200Gbit/s

**Chaque nœud :**  
2 CPU IBM Power9  
6 GPU Nvidia Tesla Volta  
2 T de memoire