

# Journée des utilisateurs Gricad

## Bilan de l'utilisation des infrastructures de calcul

GRICAD - Pôle Calcul  
Bruno Bzeznik

November 27, 2018



# Plateformes matérielles calcul GRICAD



CiGri lightweight computing grid



OAR batch scheduler

## HPC platform

*Froggy*



3200 Xeon E5 cores @2.6Ghz  
+18 GPU K20m



High performance distributed storage (Lustre): 90 TB



Infiniband FDR network



Remote visu nodes

OAR batch scheduler

## Data processing platform



*Luke*



~1000 cores – heterogeneous systems  
and continuously evolving



Local scratches on nodes  
500 TB



10 Gbe network



Remote visu nodes

## HPCDA platform

*Dahu / Yeti / Bigfoot*



2560 cores Xeon SKL Gold 6130  
192 Go RAM / node



Local SSD + NVMe



Omnipath Network  
100Gb/s

Shared with



**BETTIK: Common distributed scratch (BeeGFS) 520 To**

**MANTIS: Common distributed storage (IRODS) 1,1Po**





- ▶ **Froggy**: extension de maintenance jusqu'à mai 2019
- ▶ **Luke**: 9 nouveaux noeuds (256 coeurs)
  - ▶ -> 58 noeuds, 1162 coeurs
  - ▶ Evolution toujours hétérogène et continue (environ un nouveau noeud tous les 30/40 jours)
- ▶ **Dahu@ciment**: nouvelle plateforme 40 noeuds, 1280 coeurs, OPA, SSD, évolutive
- ▶ **Mantis** (stockage iRods) : 2 nouveaux noeuds de 100To bruts, 90To réels, chacun
- ▶ **Bettik** (scratch beeGFS) :
  - ▶ passage de 3 à 7 noeuds de data (-> 520 To)
  - ▶ remplacement du noeud de meta-data par 2 noeuds équipés de 6 SSD



- ▶ **Froggy**: néant
- ▶ **Luke**: les évolutions de Dahu seront privilégiées
- ▶ **Mantis** (stockage iRods) : fin de garantie des noeuds du projet initial - FORT BESOIN DE CONTRIBUTIONS
- ▶ **Bettik** (scratch beeGFS) :
  - ▶ Passage à réseau Omnipath
  - ▶ Ajout de noeuds en fonction des besoins (actuellement déjà plein!) - BESOIN DE CONTRIBUTIONS
- ▶ **Dahu**:
  - ▶ Ajout de 3 noeuds IA: 6 x GPU Volta V100 (en cours d'installation)
  - ▶ Extension OPA: ajout d'un switch (commandé)





- ▶ 6592 coeurs de calcul
- ▶ 169 Tflops
- ▶ Total memory: 33.616 TB
- ▶ Total disk space: 2.9 PB









## Bilan statistique Froggy + Luke

# 3 dernières années Froggy + Luke

Nombre d'heures par semaine



Jobs Count

**13,587,964**

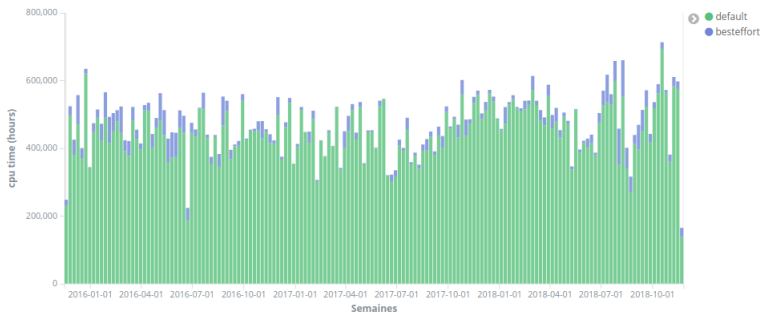
Jobs

Cpu time (h)

**74,119,531**

Hours

Temps de calcul (cpu) par semaine



# 3 dernières années Luke

Nombre d'heures par semaine



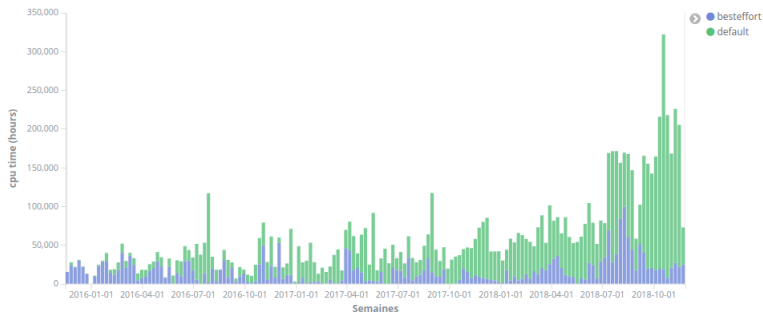
Jobs Count

**6,861,079**  
Jobs

Cpu time (h)

**9,038,606**  
Hours

Temps de calcul (cpu) par semaine



# 3 dernières années Froggy

Nombre d'heures par semaine



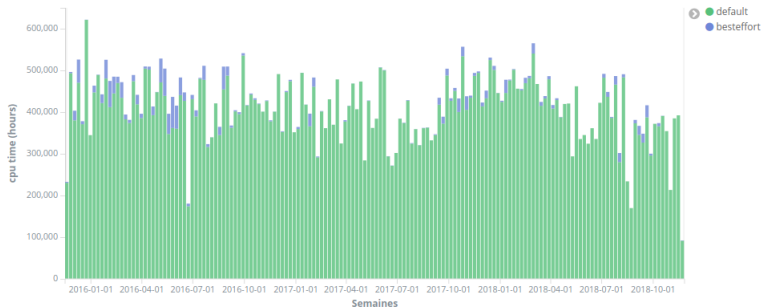
Jobs Count

**6,726,885**  
Jobs

Cpu time (h)

**65,080,925**  
Hours

Temps de calcul (cpu) par semaine





- ▶ Froggy
  - ▶ 2016: 84 %
  - ▶ 2017: 80 %
  - ▶ 2018: 80 %
- ▶ Luke
  - ▶ 2018 (de mai à mai): environ 30 %
- ▶ Dahu
  - ▶ 2018 (depuis juillet 2018) :

Année	Default	Besteffort
2016	45	4
2017	12	4
2018	29	0.5

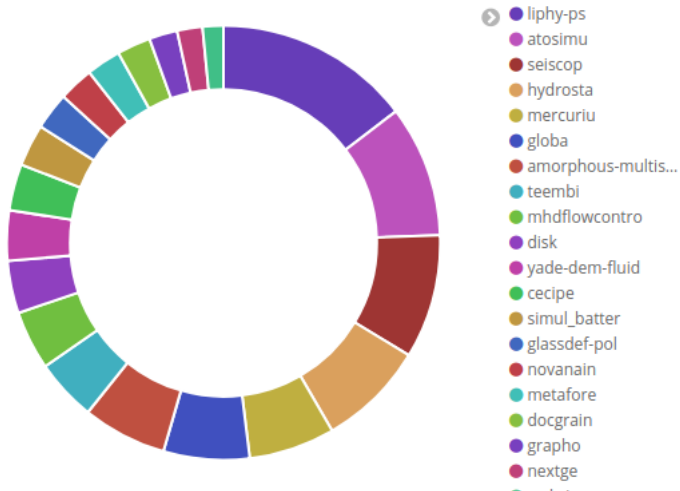
Table: Temps d'attente moyen en minutes

# Top 20 des projets - Froggy+Luke

2018



Temps cpu par projet

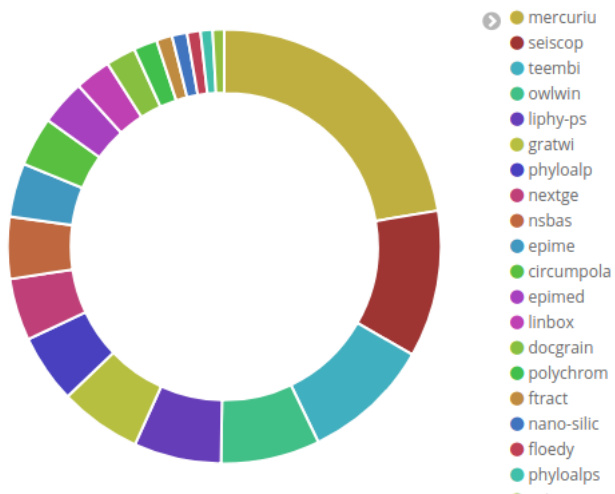


# Top 20 des projets - Luke

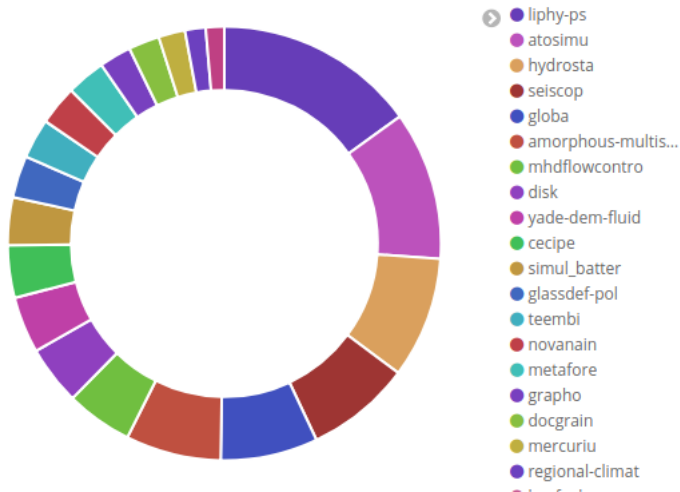
2018



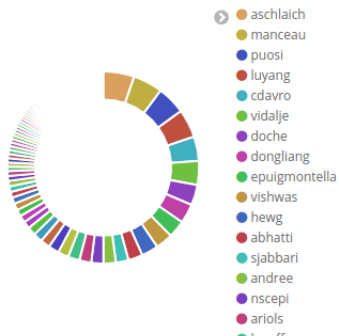
Temps cpu par projet



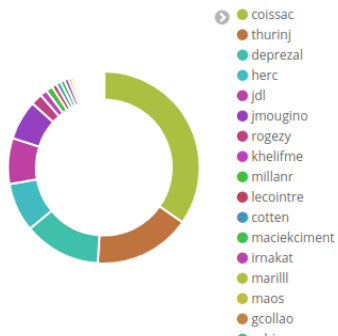
## Temps cpu par projet



Users repartition (cpu time)



Users repartition (nb jobs)

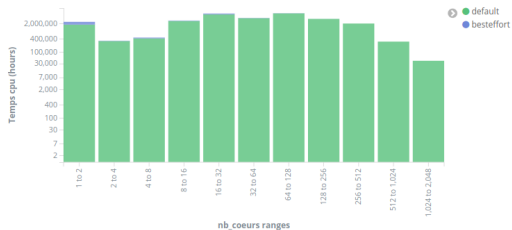


# Répartition des jobs par taille - F+L

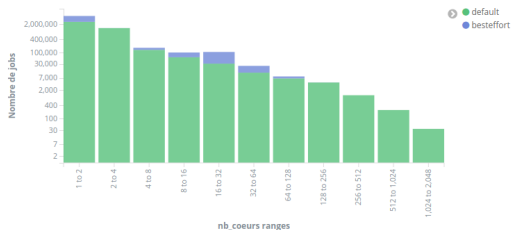
2018



Répartition des jobs par taille - temps cpu (log)



Répartition des jobs par taille - nombre (log)

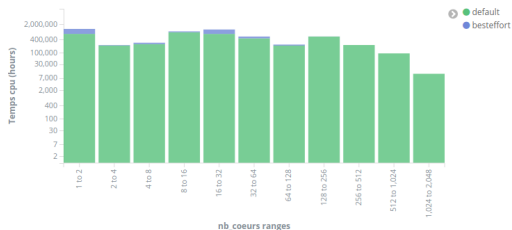


# Répartition des jobs par taille - Luke

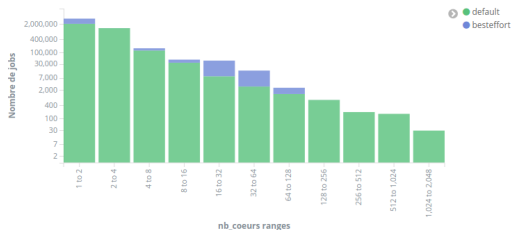
2018



Répartition des jobs par taille - temps cpu (log)



Répartition des jobs par taille - nombre (log)

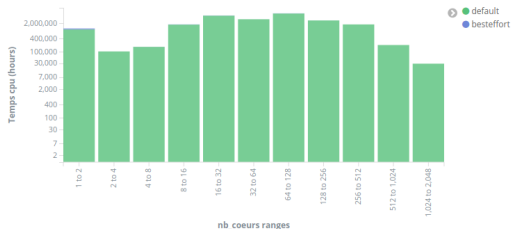




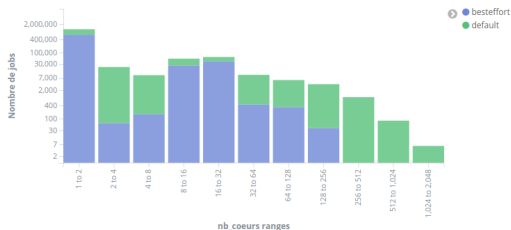
# Répartition des jobs par taille - Froggy 2018



Répartition des jobs par taille - temps cpu (log)



Répartition des jobs par taille - nombre (log)

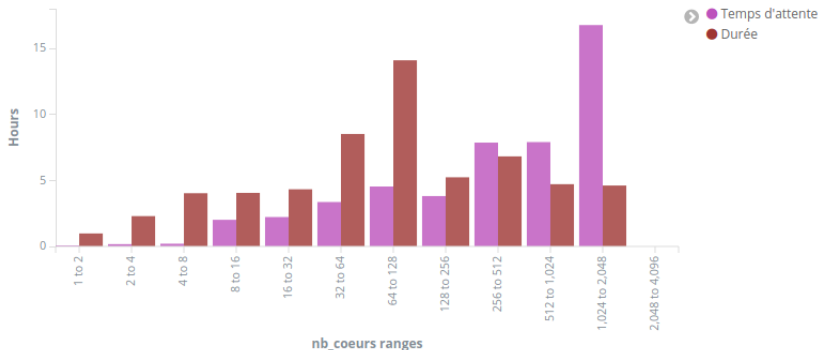


Pourcentage des heures consommées par des jobs de plus 128 coeurs:

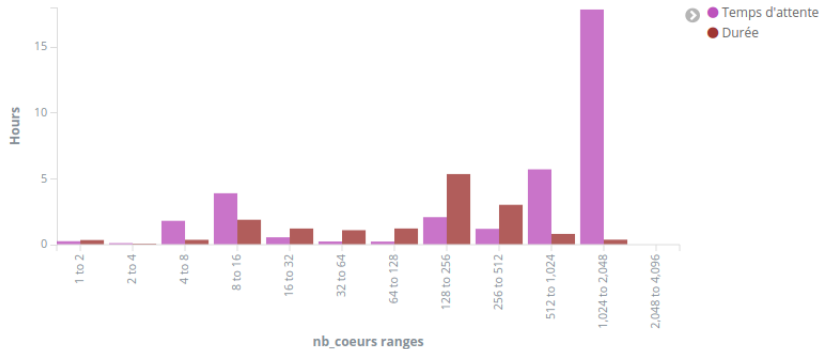
- ▶ 2016: 40%
- ▶ 2017: 27%
- ▶ 2018: 22%

Alors que le temps d'attente de ces jobs a diminué!  
Les jobs qui passent à l'échelle ont probablement migré vers les centres nationaux; cela conforte le choix de fat-tree à facteur de blocage (moins couteux) pour les futures plateformes.

## Temps d'attente et durée



Temps d'attente et durée



## Situation actuelle

- ▶ Capacité totale : 940To ; taux de remplissage : 49% ;
- ▶ Décommission récente d'un serveur défectueux et panne matérielle de deux serveurs (120To) ;
- ▶ 75% des serveurs de stockage hors garantie (340To) ;
- ▶ Difficultés d'utilisation depuis Froggy.

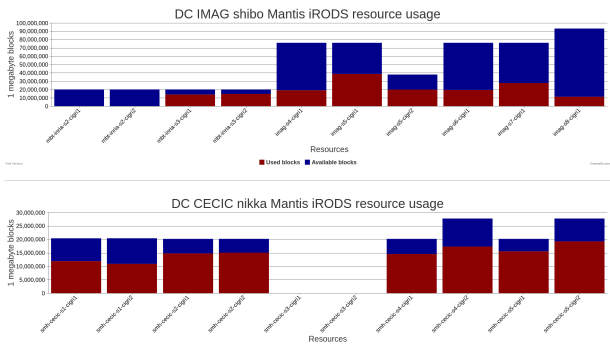
## Évolution

- ▶ Ajout d'un nouveau noeud : 90To supplémentaires à venir ;
- ▶ Décommission des serveurs hors garantie, passage de trois à deux localités (imag, simsu) ;
- ▶ Une partie des serveurs seront réinstallés avec une configuration plus résiliente aux pannes.

# Taux d'utilisation des ressources Mantis



<https://ciment-grid.ujf-grenoble.fr/mantis/mantis-resource-usage.html>

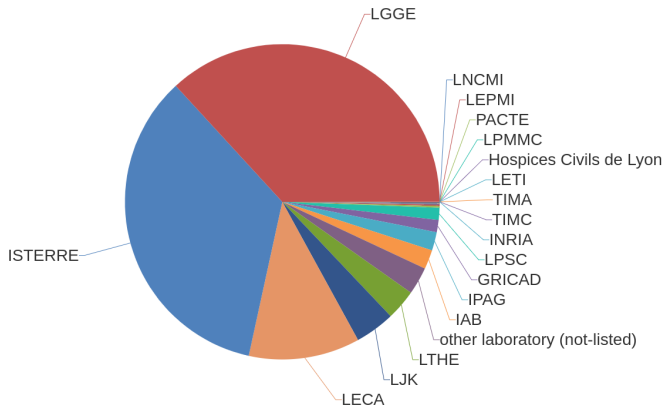


Page web actualisée toutes les heures, accessible via le serveur de la grille : <https://ciment-grid.ujf-grenoble.fr/>

# Répartition de l'utilisation de Mantis par groupe



Utilisation relative par groupe, pour un taux d'utilisation global de 49%



## Questions