

Étude des applications *Bag-of-Tasks*  
du méso-centre Gricad  
COMPAS22

Quentin GUILLOTEAU, Olivier RICHARD, Eric RUTTEN

Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG

2022-07-07

# Le méso-centre *Gricad*

## *Gricad*

- Méso-centre @Grenoble
- Calculs Scientifiques
- 5 Grappes de calculs (HPC, HPCDA, ...)



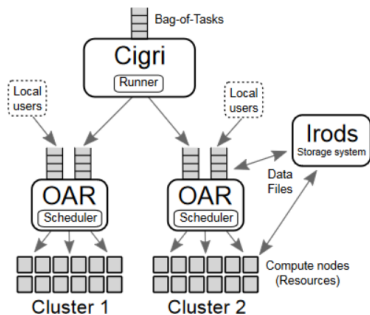
## Un problème rencontré

Ressources HPC inutilisées  $\implies$  Perte de puissance de calcul  
 $\hookrightarrow$  **Comment récolter ?**

# L'intergiciel *CiGri*

## Une solution: *CiGri*

- Intergiciel de Grille
- **bag-of-tasks**: Bcp, multi-param
- **Tâches Best-effort**: Prio la + basse
- Réduit la pression sur le RJMS
- **Objectif**: Collecter les ressources libres de la grille



## Applications *Bag-of-Tasks*

- Simulations type Monte-Carlo
- Parameter Sweep

## Soumissions *CiGri*

Tâche ∈ **Campagne** ∈ Projet

## Un exemple de projet: Biggnss

- **Big Data + Global Navigation Satellite System**
- Traitement massif de données GPS pour le traitement de la déformation de la surface terrestre
- Bcp satellites  $\implies$  bcp de données
- Plusieurs *stations* géographiques  $\rightsquigarrow$  **Campagnes**
- Subdivision du traitement  $\rightsquigarrow$  **Tâches**
- Unique binaire + différents paramètres

# Les limites de *CiGri*

## Quel est le besoin de cette étude ?

### Le problème de *CiGri*

↗ Collecte  $\implies$  ↗ Perturbations (e.g., I/O)  $\rightsquigarrow$  **Trade-off**

$\leftrightarrow$  Imprévisibilité  $\implies$  **régulation en ligne**

### Notre approche (présentée @COMPAS21!)

Utiliser la **Théorie du Contrôle** pour réguler les soumissions de *CiGri*

### Notre but

Augmenter l'**utilisation des grappes** et **garantir une QoS**

# CiGri + Mécanisme de rétro-action

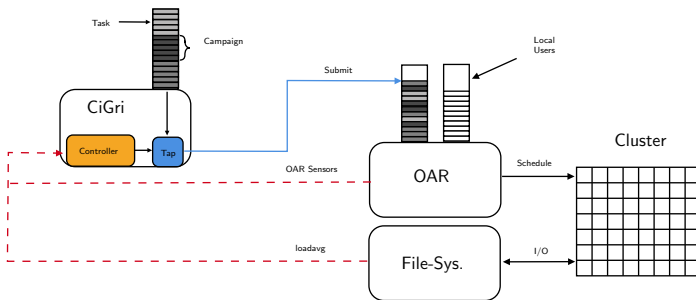
## Objectif

Collecter les ressources libres  
**sans perturbations**

- max utilisation grappes
- min perturbations

## Comment?

- Instrumentation
  - **Actionneur**: #tâches, ...
  - **Capteurs**: RJMS WQ, charge FS
- **Controlleurs** (PID, RST, MFC, ...)
- **Validation expérimentale**



# Setup expérimental

## Pas possible de faire des expériences à taille réelle

### Expériences = Emulation

- Grid'5000
- **Repliment**: 1 grappe sur une machine
- Pas de "vraies" tâches
- Tâches = `sleep` + I/O
- CPU Bound + I/O Bound
- (Simulation  $\rightsquigarrow$  même problème)

↪ Comment représenter de **manière réaliste** les campagnes BoT ?

## Objectifs de cette étude

### Questions au niveau global

- Distribution des durées des tâches BoT ?
- Quantité de tâches BoT ?
- Evolution des métriques au cours du temps ?

### Questions au niveau des projets

- Combien de tâches dans une projet ?
- Modélisation des durées ?

### Question au niveau des campagnes

- **Combien de tâches par campagne ?**
- **Distribution des durées des tâches dans une campagne ?**



1 Introduction, Motivation & Concepts

2 Questions au niveau global

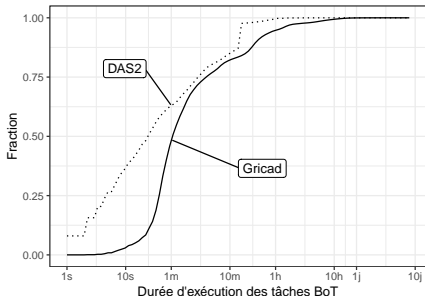
3 Questions au niveau des projets

4 Questions au niveau des campagnes

5 Conclusion & Perspectives

## Distribution des durée des tâches

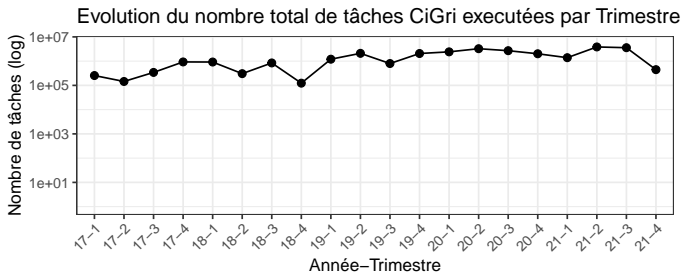
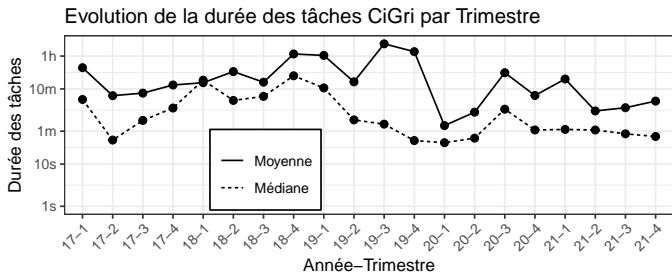
Fonctions de répartition des durées des tâches BoT pour les grilles Gricad et DAS2



Métriques	Grilles de Calculs	
	<i>Gricad</i>	<i>DAS2</i>
Nombre de tâches BoT	$\simeq 3 \times 10^7$	$\simeq 10^5$
Nombre de grappes	5	5
Minimum $t_{exec}$	1s	1s
Moyenne $t_{exec}$	19m	3m 51s
Maximum $t_{exec}$	7j 22h 20m	10h
Médiane $t_{exec}$	1m 3s	24s
Quantile 75 % $t_{exec}$	3m 55s	2m 51s
Quantile 95 % $t_{exec}$	1h 5m 38s	15m 1s
Quantile 99 % $t_{exec}$	7h 28m 40s	43m 35s

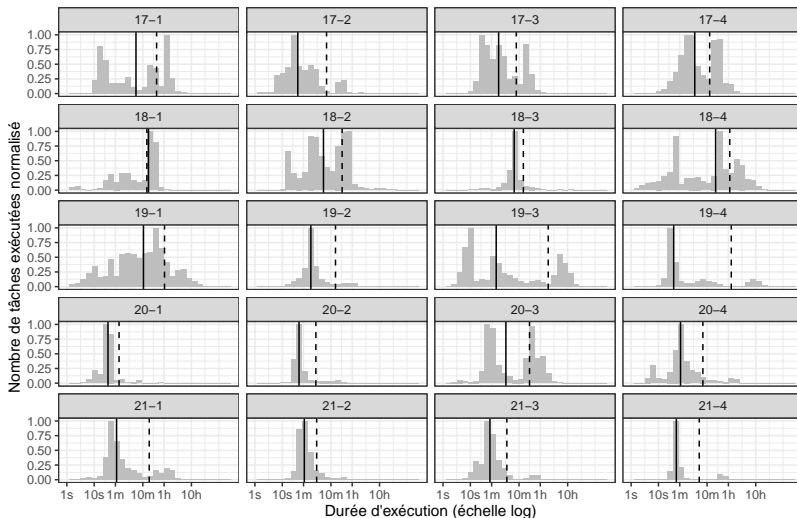
↔ 50% < 1min & 75% < 4mins ⇒ beaucoup de **petites tâches**

# Evolution des métriques globales



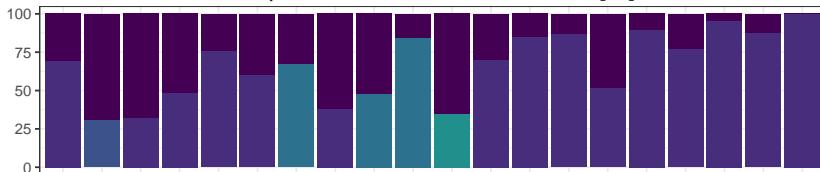
# Evolution de la distribution durées des tâches

Distribution des durées d'exécution des tâches CiGri par Trimestre

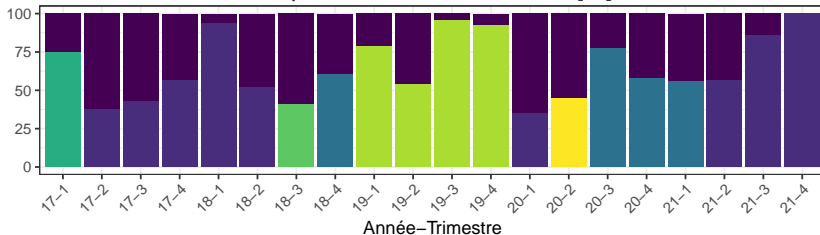


# Proportion des tâches exécutées et travail par projet

## Proportion des tâches exécutées [%]



## Proportion du travail effectué [%]



1 Introduction, Motivation & Concepts

2 Questions au niveau global

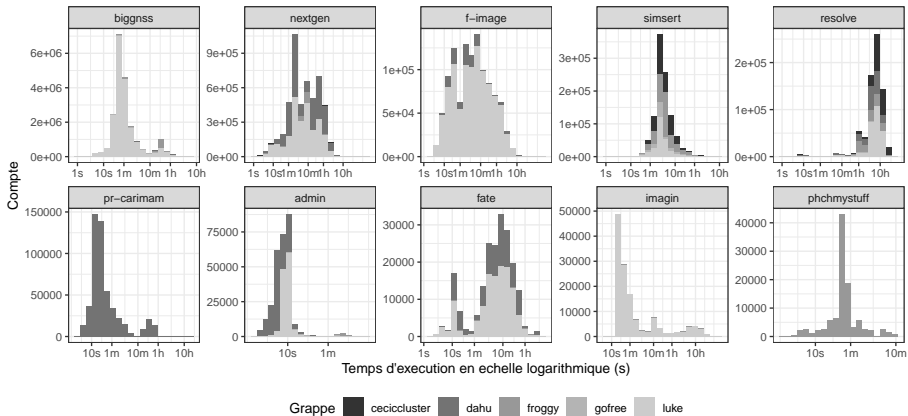
3 Questions au niveau des projets

4 Questions au niveau des campagnes

5 Conclusion & Perspectives

# Distribution des durées d'exécution par projet

Histogramme des temps d'exécution des tâches des 10 projets avec le plus de tâches



↪ **Différentes distributions par projet et par grappe**

# Modélisation des distributions

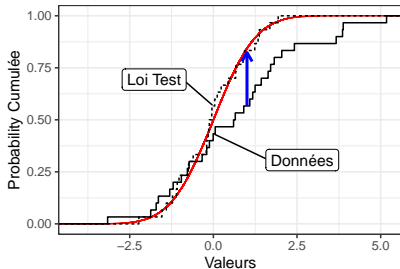
## Objectif

**Trouver la loi de probabilité des durées d'exécution**

## Protocole

- Fit des distributions sur les données
- Test de fit Cramér-Von Mises
- 30 répétitions

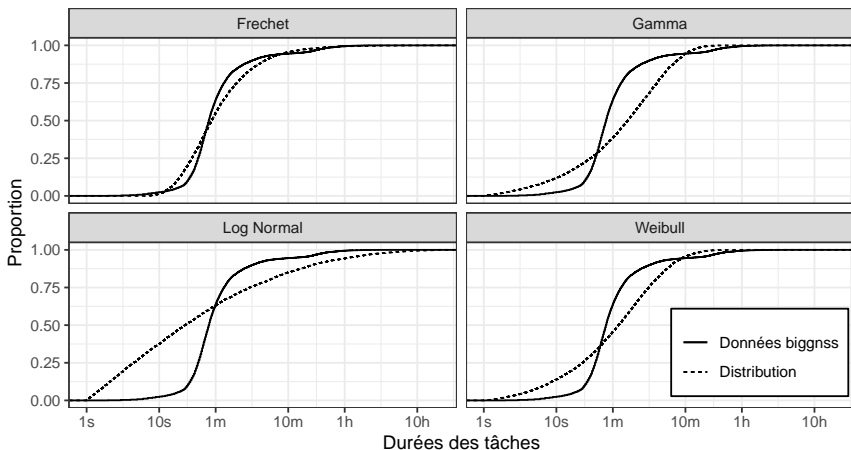
Principe du test Cramér-von Mises



## Distributions à queues lourdes

↪ Log-Normale, Weibull, Frechet, Gamma (+ Normale)



Exemple sur le projet *biggnss*Ajustement de plusieurs distributions sur les données de *biggnss* sur Luke

↪ La distribution *Frechet* semble le plus s'ajuster aux données

Projet	Grappe	Distribution
admin	Dahu	<i>Log-Normale</i>
	Froggy	<i>Frechet</i>
	Gofree	<i>Gamma</i>
	Luke	<i>Log-Normale</i>
biggnss	Dahu	<i>Normale</i>
	Froggy	<b>X</b>
	Gofree	<b>X</b>
	Luke	<i>Frechet</i>
f-image	Dahu	<i>Frechet</i>
	Luke	<i>Log-Normale</i>
fate	Dahu	<i>Weibull</i>
	Luke	<i>Weibull</i>
imagin	Froggy	<i>Normale</i>
	Luke	<i>Weibull</i>

Projet	Grappe	Distribution
nextgen	Ceci	<i>Frechet</i>
	Dahu	<i>Weibull</i>
	Froggy	<i>Weibull</i>
	Luke	<i>Gamma</i>
phcmystuff	Froggy	<i>Log-Normale</i>
pr-carimam	Dahu	<i>Frechet</i>
resolve	Ceci	<i>Gamma</i>
	Dahu	<i>Normale</i>
	Froggy	<i>Normale</i>
	Luke	<i>Normale</i>
simsert	Ceci	<i>Frechet</i>
	Dahu	<i>Frechet</i>
	Froggy	<i>Frechet</i>
	Luke	<i>Frechet</i>

↪ **En accord avec d'autres workloads de grilles [losup,+08]**

1 Introduction, Motivation & Concepts

2 Questions au niveau global

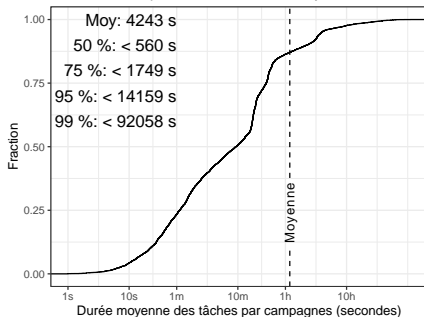
3 Questions au niveau des projets

4 Questions au niveau des campagnes

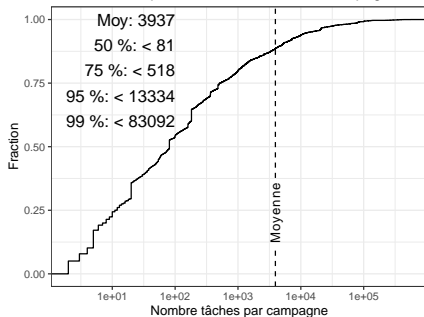
5 Conclusion & Perspectives

# Caractéristiques par Campagne

Fonction de répartition de la durée moy. des tâches

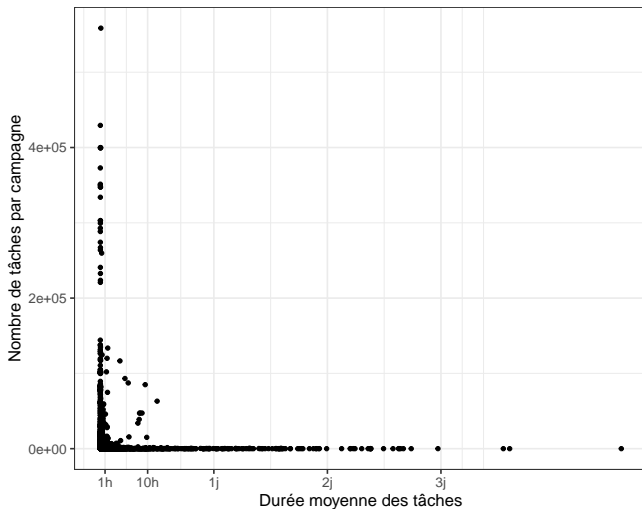


Fonction de répartition du nb de tâches/campagne



↔ Durée moyenne pour des petites campagnes ? pour des grandes ?

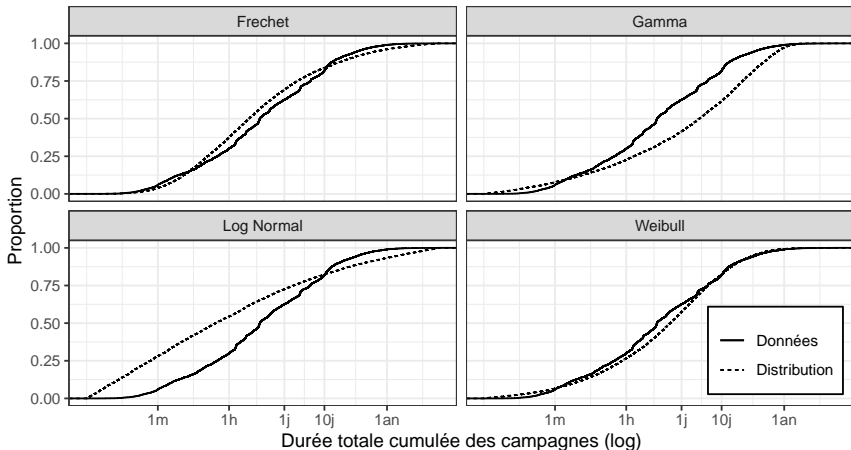
## Relation entre durée moyenne et nombre de tâches par campagne



↪ Beaucoup de grandes campagnes avec des tâches courtes  
& des petites campagnes avec des tâches longues

# Modélisation du travail total par campagne

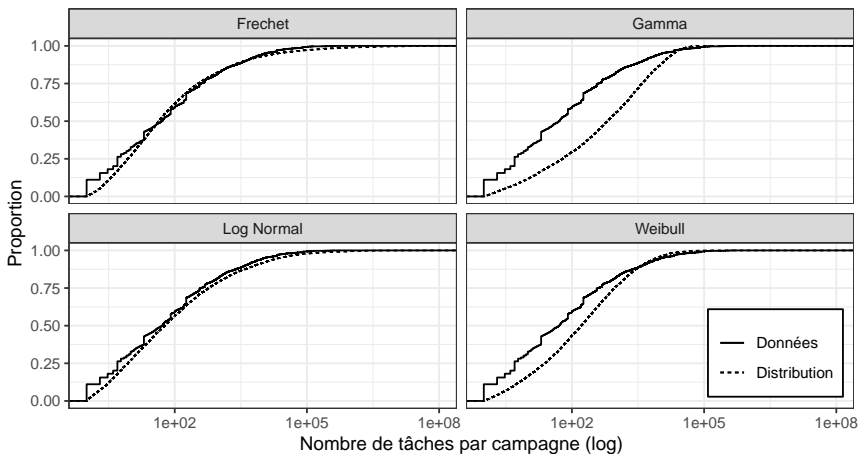
Ajustement de plusieurs distributions sur les durées cumulées par campagne



↪ La distribution *Weibull* semble le plus s'ajuster aux données

# Modélisation du nombre de tâches par campagne

Ajustement de plusieurs distributions sur le nb de tâches par campagne



↪ **Les distributions *Frechet* & *Log-Normal* semblent les plus s'ajuster aux données**

## Comment (grossièrement) générer une campagne ?

- 1 Tirer une quantité de travail
- 2 Tirer un nombre de tâches dans la campagne
- 3 Durée moyenne des tâches = travail / nb de tâches
- 4 Choix de la distributions des durées dans la campagne



1 Introduction, Motivation & Concepts

2 Questions au niveau global

3 Questions au niveau des projets

4 Questions au niveau des campagnes

5 Conclusion & Perspectives

# Conclusion & Perspectives

## Rappel de l'objectif

**Etudier les applications BoT de Gricad + essayer de les modéliser**

## Conclusion

- Métriques globales (durées, nb tâches) + évolution
- Distribution des durées par projet + modélisation
- Métriques par campagne (durées, nb tâches) + modélisation

## Perspectives

- Pourquoi queues lourdes ?
- Motifs I/O ?
- Grouper les campagnes ?
- Vos idées/questions (!)

↔ Données + Scripts: <https://zenodo.org/record/6787030>