

CiGri version 3

Journée des utilisateurs CIMENT

Bruno Bzeznik

2014-05-14



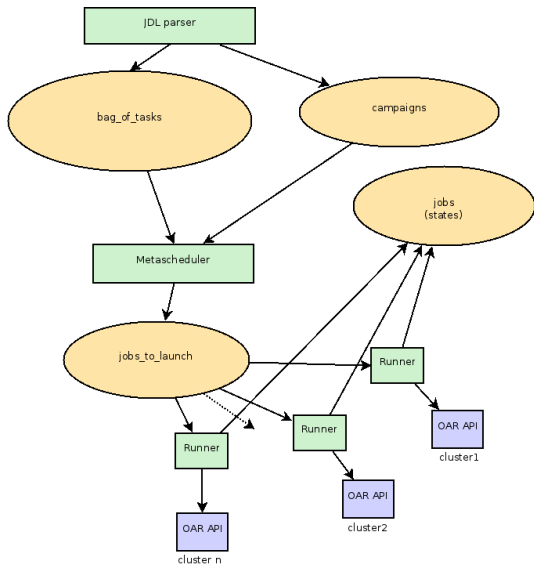
- Projet du début de CIMENT, vers 2000.
- Exploiter les CPU libres des différentes machines de calcul du mesocentre avec des jobs *embarassingly parallel* (de gros sacs de tâches indépendantes)
- Le code (perl) est développé au sein de CIMENT et en collaboration avec l'équipe MESCAL du LIG. C'est toujours le cas aujourd'hui.
- Jusqu'en 2006, le développeur principal est Nicolas Capit
- A partir de 2006 le code est repris par Bruno Bzeznik
- En 2010, Ciment met en place un système de stockage distribué (IRODS) qui complètera efficacement CiGri pour tous les aspects gestion des données. Certaines parties du code pourront être supprimées grace à cette évolution
- En 2012/2013, CiGri est complètement ré-écrit, en Ruby en collaboration avec Grid5000 (le développeur principal est toujours Bruno Bzeznik, aidé par un développeur de Grid5000, Ghislain Charier). Certains concepts sont complètement revus, à la suite d'un travail au sein de MESCAL et de GRID5000.
- Fin 2013, CiGri v3 est en production sur CIMENT.
- En mai 2014, CiGri v3 a déjà exécuté près de 4 millions de jobs

- OAR sur tous les clusters de CIMENT
- Un job best-effort à une priorité nulle
- Un job best-effort peut être tué à n'importe quel moment si un autre job a besoin des ressources
- CiGri permet d'exploiter efficacement ce type de jobs en gérant la re-soumission automatique
- Mais CiGri v3 peut aussi maintenant gérer des jobs normaux (non best-effort)

Principales nouveautés

- Code évolutif en Ruby et PostgreSQL
- Nouveaux fichiers JDL en JSON
- Communication par API REST (client et serveur)
- Affinités Users/Clusters
- Jobs prologue/epilogue
- Système de notifications évolué (avec niveaux, groupage, choix des destinataires,...)
- Gestion des évènements évoluées (reprise du stderr, ...)
- Contrôle complet par la ligne de commande (ou API)
- Overhead de soumission grandement réduit (gestion des array jobs)
- Détection des surcharges
- Définition des ressources "à la OAR"
- Test mode
- "max jobs" limit

Principe



Exemple de fichier JDL

```
{
  "name": "test_prog",
  "resources": "core=1",
  "exec_file": "$HOME/launch.bash",
  "exec_directory": "$HOME",
  "param_file": "absolute path to your file parameter/param.txt",
  "test_mode": "false",
  "type": "best-effort",
  "clusters": {
    "gofree": {
      "max_jobs": "10",
      "project": "test",
      "walltime": "00:05:00"
    },
    "froggy": {
      "max_jobs": "10",
      "project": "test",
      "walltime": "00:05:00"
    }
  },
  "prologue": [
    "set -e",
    "source /applis/ciment/v2/env.bash",
    "module load irods",
    "cd $HOME",
    "secure_iget -f launch.bash",
    "chmod 777 launch.bash",
    "secure_iget -f prog.py",
    "chmod 777 prog.py"
  ]
}
```

Soumission : gridsub

```
daemon@killeen:~$ gridsub -f povray.jdl  
Campaign successfully submitted  
CAMPAIGN_ID=1415
```

gridstat

```
daemon@killeen:~$ gridstat
```

Campaign id	Name	User	Submission time	S	Progress
979	test_prog	vcvargas	2014-02-27 11-47-44	Re	3/4 (75%)
1085	dzero_pmcs007_extra	biscarat	2014-03-18 16-52-52	Re	0/37 (0%)
1328	AS_MODECOGEL	viry	2014-04-22 17-50-10	Re	4753/5000 (95%)
1383	RecupResultsBootstr	dgeorges	2014-04-29 18-51-14	Re	0/100 (0%)
1408	no_simsus4	lecointre	2014-05-05 07-42-23	Re	319/351 (90%)
1410	pVM_2010	lecointre	2014-05-06 10-46-25	R	253/365 (69%)
1413	test_V3CorrelationT	briandx	2014-05-07 13-55-22	R	4106/6710 (61%)
1414	BootstrapALLSndRun	dgeorges	2014-05-07 14-18-00	R	0/4080 (0%)

gridstat

```
daemon@killeen:~$ gridstat 1413
Campaign: 1413
Name: test_V3CorrelationTraitement
User: briandx
Date: 2014-05-07 13-55-22
State: in_treatment
Progress: 4357/6710 (64%)
Stats:
  average_jobs_duration: 249.088707949432
  stddev_jobs_duration: 227.454207425469
  jobs_throughput: ~ jobs/h
  remaining_time: 0.6 hours
  failures_rate: 7.4 %
  resubmit_rate: 7.4 %
Clusters:
  froggy:
    active_jobs: 143
    queued_jobs: 89
    prologue_ok: true
    epilogue_ok: true
  luke:
    active_jobs: 18
    queued_jobs: 12
    prologue_ok: true
    epilogue_ok: true
```

gridevents

```
daemon@killeen:~$ gridevents -c 1408
```

```
-----  
1758979: (open) EXIT_ERROR of job 3763874 at 2014-05-07T14:56:05+02:00 on froggy
```

```
The job exited with exit status 2;
```

```
Last 5 lines of stderr_file:
```

```
ERROR: connectToRhost: error returned from host 152.77.100.4 status = -118000 status = -118000 SYS_MAX_CO
```

```
ERROR: _rcConnect: connectToRhost error, server on 152.77.100.4 is probably down status = -118000 SYS_MAX
```

```
-----  
1758980: (open) BLACKLIST at 2014-05-07T14:56:05+02:00 on froggy because of 1758979
```

```
-----  
daemon@killeen:~$ gridevents -f -r -e 1758979
```

gridclusters

```
deamon@killeen:~$ gridclusters -b
1: gofree , gofree.ujf-grenoble.fr (stress:0.11/0.8) (336 resources)
2: froggy , froggy.ujf-grenoble.fr (stress:0.19/0.8) (3268 resources)
3: fontaine , fontaine.ujf-grenoble.fr (stress:0/0.8) (172 resources)
4: fostino , fostino.obs.ujf-grenoble.fr (stress:0.14/0.8) (464 resources)
5: luke , luke.ujf-grenoble.fr (stress:0.62/0.8) (108 resources)

- =unavailable - =used - =used_by_cigri - =free
deamon@killeen:~$
```

gridnotify

```
bzizou@killeen:~$ gridnotify -l
```

You have the following notification subscriptions:

- mail on Bruno.Beznik@imag.fr with severity medium
- jabber on bzizounet@gmail.com with severity low

Stockage distribué : IRODS

- Ciment fourni un stockage distribué qui permet d'échanger des données de manière uniforme sur tous les clusters, avec un espace de noms unique (environ 700TB sous Irods)
- Pas de lien direct entre Cigri et Irods
- Irods apparait rapidement complémentaire et indispensable à Cigri : le lien se fait dans les prologues (diffusion du code et scripts de calcul) et les jobs eux-même (i/o de calcul)
- Pre-configuration de l'authentification Irods sur les clusters pour que l'utilisation soit immédiate et transparente

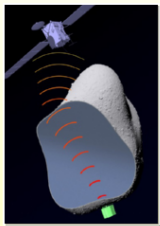
- Un portail web
- Des stats d'utilisation
- Une doc plus étoffée
- Des ajouts de fonctionnalités divers (macros variables dans le JDL, gridclean,...)

Exemples d'utilisation de Cigri v3



ROSETTA / Comet Nucleus Sounding Experiment

by Radiowave Transmission



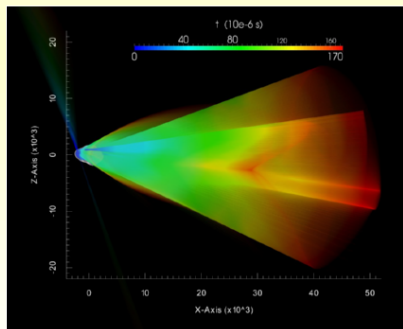
Dans le cadre de la mission spatiale de l'ESA : ROSETTA,

CONCERT effectue une tomographie radar du noyau de la comète 67P/Churyumov-Gerasimenko.

2014 : La grille CIMENT et Cigri v3 sont utilisés pour la préparation des opérations spatiales de l'instrument CONCERT:

- Identification des fenêtres temporelles permettant des mesures efficaces en fonction des paramètres orbitaux.
- Simulation numérique de l'expérience.

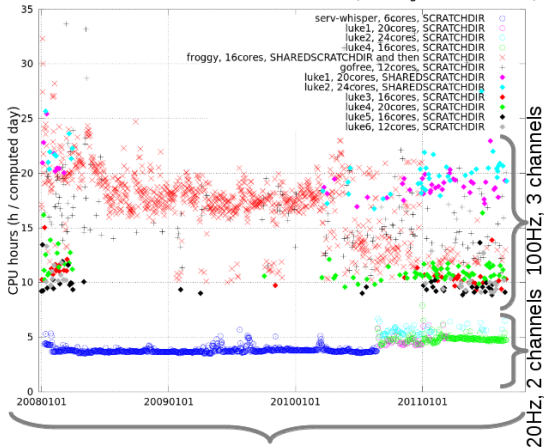
2015 : exploitation des données d'observation pour la caractérisation du noyau cométaire



réseau HiNet
~800stations

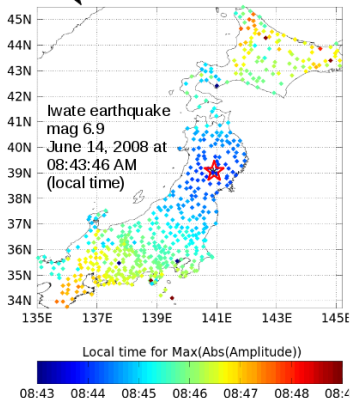


CPU cost for the CNT -> MSED Float32 conversion (including IO iRODS transfer)



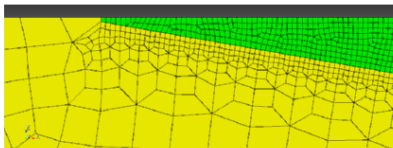
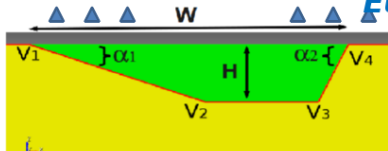
2 x 1339 jobs parallèles (1job=1node) CiGri v3

VM SHZ 20080614 T<5s

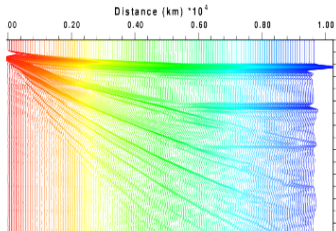


Seismic amplification in alluvial valleys

European project NERA



Automated FE mesh design



Goal: Identify the parameters that control the amplification of earthquake ground motion in sedimentary basins or alluvial valleys. Account for 2D/3D effects in seismic design codes.

2D Trapezoidal or triangular valleys

Control parameters: H, W, α_1, α_2

162 valleys * 6 velocity models * 2 excitations =

1944 simulations on 8-96 cores

Submissions through cigri3

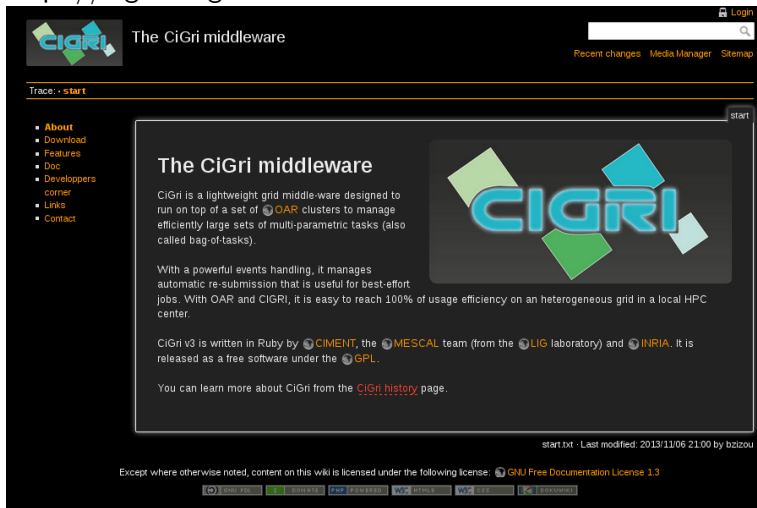
280 000 1C-hours on Froggy

About 1 million seismic traces stored on IRODS (56 Gbytes only) with user metadata

Post-processing will generate ~10 times more traces

Test of IRODS python

http ://cigri.imag.fr



The screenshot shows the homepage of the CiGri middleware project. At the top left is the CiGri logo, followed by the text "The CiGri middleware". On the right, there is a search bar, a "Login" link, and links for "Recent changes", "Media Manager", and "Sitemap". Below the header is a navigation menu with "Trace" and "start" options. A sidebar on the left contains a list of links: "About", "Download", "Features", "Doc", "Developpers corner", "Links", and "Contact". The main content area features a large heading "The CiGri middleware" next to a large CiGri logo. Below the heading is a paragraph describing CiGri as a lightweight grid middle-ware designed to run on top of a set of OAR clusters to manage efficiently large sets of multi-parametric tasks (also called bag-of-tasks). Another paragraph explains that with a powerful events handling, it manages automatic re-submission that is useful for best-effort jobs. With OAR and CiGri, it is easy to reach 100% of usage efficiency on a heterogeneous grid in a local HPC center. A third paragraph states that CiGri v3 is written in Ruby by the CIMENT, the MESCAL team (from the LIG laboratory) and INRIA, and is released as a free software under the GPL. A link to the "CiGri history" page is provided. At the bottom of the main content area, it says "start.txt · Last modified: 2013/11/06 21:00 by bziouz". At the very bottom, there is a license notice: "Except where otherwise noted, content on this wiki is licensed under the following license: GNU Free Documentation License 1.3" and a row of icons for various actions like "edit", "create", "delete", "help", "info", "login", and "register".

The CiGri middleware

Trace: · start

- About
- Download
- Features
- Doc
- Developpers corner
- Links
- Contact

The CiGri middleware

CiGri is a lightweight grid middle-ware designed to run on top of a set of [OAR](#) clusters to manage efficiently large sets of multi-parametric tasks (also called bag-of-tasks).

With a powerful events handling, it manages automatic re-submission that is useful for best-effort jobs. With OAR and CiGri, it is easy to reach 100% of usage efficiency on a heterogeneous grid in a local HPC center.

CiGri v3 is written in Ruby by [CIMENT](#), the [MESCAL](#) team (from the [LIG](#) laboratory) and [INRIA](#). It is released as a free software under the [GPL](#).

You can learn more about CiGri from the [CiGri history](#) page.

start.txt · Last modified: 2013/11/06 21:00 by bziouz

Except where otherwise noted, content on this wiki is licensed under the following license: [GNU Free Documentation License 1.3](#)

[edit](#) [create](#) [delete](#) [help](#) [info](#) [login](#) [register](#)