



GREASY

Paralelismo trivial en LaPalma

SIMULACIONES HPC EN LAPALMA

- Cada subproblema interactúa con alguno o varios subproblemas
- Necesita una red de interconexión de alto rendimiento
- Ejecución en LaPalma

```
#!/bin/bash
```

```
#
```

```
# @ job_name = simul01
```

```
# @ initialdir = .
```

```
# @ output = simul_%j.out
```

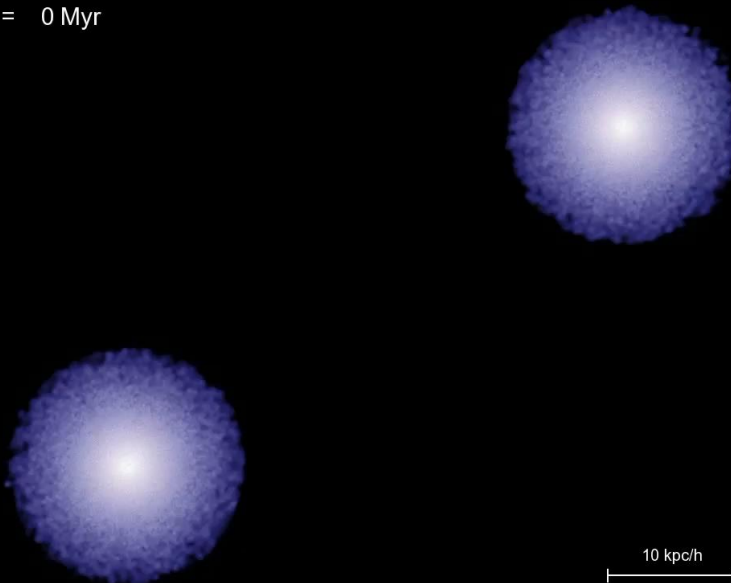
```
# @ error = simul_%j.err
```

```
# @ total_tasks = 64
```

```
# @ wall_clock_limit = 72:00:00
```

```
srun ./Gadget2 simul01.tex
```

T = 0 Myr



SIMULACIONES HPC EN LAPALMA (II)

- Lista de trabajos en el sistema de gestión de colas

```
~/galaxy> mnq
```

```
active jobs-----
```

NAME	JOBID	STATE	PROCS
simu101	2010242	Running	64

PARALELISMO TRIVIAL (I)



- Tratamiento de un gran número de imágenes
- Pre y postprocesamiento de código paralelo
- Ajuste de parámetros en código serial
- Ejecución en LaPalma

```
#!/bin/bash
#
# @ job_name = recognition
# @ initialdir = .
# @ output = recognition_%j.out
# @ error = recognition_%j.err
# @ total_tasks = 1
# @ wall_clock_limit = 00:05:00
./face_recog img01.jpg
```

PARALELISMO TRIVIAL (II)

- Escribir un fichero de comandos por cada trabajo de 1 procesador
- Tamaño considerable de la lista de trabajos

```
~/face_recog> mnq
```

```
active jobs-----
```

NAME	JOBID	STATE	PROCS
recognition	2010381	Running	1
recognition	2010382	Running	1
recognition	2010383	Running	1
recognition	2010384	Running	1
recognition	2010385	Running	1
recognition	2010386	Running	1
recognition	2010387	Running	1

```
...
```

GREASY

- Herramienta desarrollada por el BSC y disponible en todas las máquinas de la RES.
- Simplifica el paralelismo trivial agrupando múltiples tareas en un único trabajo.
- En un sistema como LaPalma la cola de trabajos se reduce considerablemente.
- 1 fichero, múltiples líneas, cada línea 1 tarea.
- Único fichero de registro para el conjunto de tareas.
- Posibilidad de indicar dependencias.

SOFTWARE GREASY

- Se puede encontrar en `/gpfs/apps/GREASY` en todas las máquinas de la RES

- `/gpfs/apps/GREASY/latest`

- `/bin` → Programa para invocar a greasy

- `/etc` → Configuración global

- `/doc` → Guía de usuario

- `/example` → Ficheros de ejemplo para comenzar a utilizar greasy

PRIMEROS PASOS GREASY (I)

- Copiar ficheros de ejemplo del directorio /example

`bsc_greasy.job`

`example.txt`

`short-example.txt`

- Listado de tareas en `short-example.txt`:

`/bin/sleep 20`

`/bin/sleep 10`

`/bin/sleep 5`

PRIMEROS PASOS GREASY (II)

- Fichero de comandos "bsc_greasy.job":

```
...
# @ total_tasks = 4
...
# @ wall_clock_limit = 00:05:00
...
FILE=short-example.txt
...
#####
# Run greasy!                                     #
#####
/gpfs/apps/GREASY/2.1/bin/greasy $FILE
```

- Enviar trabajo a sistema de colas (SLURM)

```
~/greasy> msubmit bsc_greasy.job
Submitted batch job 2010356
```

PRIMEROS PASOS GREASY (III)

- Trabajos en cola

```
~/greasy> mnq
```

```
active jobs-----
```

NAME	JOBID	STATE	PROCS
greasy	2010356	Running	4

```
1 active jobs
```

- Ficheros de salida

```
greasy-2010356.err
```

```
greasy-2010356.log
```

```
greasy-2010356.out
```

PRIMEROS PASOS GREASY (IV)

- Fichero de registro

```
~/greasy> cat greasy-2010356.log | greasycolorlog
Start greasing ~/greasy/short-example.txt
INFO: MPI engine is ready to run with 3 workers
INFO: Allocating task 1
INFO: Allocating task 2
INFO: Allocating task 3
INFO: Task 3 completed successfully on node s04c3b08. Elapsed: 00:00:05
INFO: Task 2 completed successfully on node s04c3b08. Elapsed: 00:00:10
INFO: Task 1 completed successfully on node s04c3b08. Elapsed: 00:00:20
INFO: MPI engine finished
INFO: Summary of 3 tasks: 3 OK, 0 FAILED, 0 CANCELLED, 0 INVALID.
INFO: Total time: 00:00:20
INFO: Resource Utilization: 58.33%
Finished greasing ~/greasy/short-example.txt
```

FICHERO DE TAREAS

- El número de línea es el identificador de una tarea
- Se pueden incluir comentarios iniciando la línea con #
- Las líneas en blanco y los comentarios también “consumen” un identificador de tarea
- Se pueden indicar dependencias incluyendo al inicio de la línea:
[# <lista de dependencias> #]
 - Sólo se permiten dependencias de tareas anteriores, no posteriores
 - Posibilidad de incluir rangos así como dependencias relativas.

DEPENDENCIAS

- Ejemplo de dependencias “example.txt”

```
/bin/sleep 1
/usr/bin/hostname
/bin/sleep 3
/bin/sleep 4
/bin/sleep 5
/bin/sleep 6
```

```
#####
# These Lines with dependencies are correct
#####
[# 5 #]          /bin/sleep 11
[# -1#]          /bin/sleep 12
[#1-3#]          /bin/sleep 13
[#1, 3 - 5 #]   /bin/sleep 14
[##]             /usr/bin/hostname
```

DEPENDENCIAS (II)

```
~/greasy> cat greasy-2010379.log | greasycolorlog
Start greasing ~/greasy/example.txt
INFO: MPI engine is ready to run with 3 workers
INFO: Allocating task 1
INFO: Allocating task 2
INFO: Allocating task 3
ERROR: Task 2 failed with exit code 32512 on node s04c4b03. Elapsed:
00:00:00
WARNING: Cancelling task 13 because of task 2 failure
INFO: Allocating task 4
INFO: Task 1 completed successfully on node s04c4b03. Elapsed: 00:00:01
INFO: Allocating task 5
INFO: Task 3 completed successfully on node s04c4b03. Elapsed: 00:00:03
INFO: Allocating task 6
INFO: Task 4 completed successfully on node s04c4b03. Elapsed: 00:00:04
INFO: Allocating task 15
ERROR: Task 15 failed with exit code 32512 on node s04c4b03. Elapsed:
00:00:00
```

DEPENDENCIAS (III)

```
INFO: Task 5 completed successfully on node s04c4b03. Elapsed: 00:00:05
INFO: Allocating task 11
INFO: Allocating task 14
INFO: Task 6 completed successfully on node s04c4b03. Elapsed: 00:00:06
INFO: Task 11 completed successfully on node s04c4b03. Elapsed: 00:00:11
INFO: Allocating task 12
INFO: Task 14 completed successfully on node s04c4b03. Elapsed: 00:00:14
INFO: Task 12 completed successfully on node s04c4b03. Elapsed: 00:00:12
INFO: MPI engine finished
INFO: Summary of 11 tasks: 8 OK, 2 FAILED, 1 CANCELLED, 0 INVALID.
INFO: Total time: 00:00:29
INFO: Resource Utilization: 64.36%
INFO: Creating restart file ~/greasy/example.txt.rst...
INFO: Restart file created
Finished greasing ~/greasy/example.txt
```

FICHERO DE REINICIO (I)

- Permite relanzar aquellas tareas erróneas

```
#  
# Greasy restart file generated at 2013-05-27 20:14:40  
# Original task file: example.txt  
# Log file: ~/greasy/greasy-2010379.log  
#  
  
# Warning: Task 2 failed  
/usr/bin/hostname  
# Warning: Task 13 was cancelled due to a dependency failure  
[# 8 #] /bin/sleep 13  
# Warning: Task 15 failed  
/usr/bin/hostname  
  
# End of restart file
```


FICHERO DE REINICIO (I)

- Permite relanzar aquellas tareas erróneas

```
#  
# Greasy restart file generated at 2013-05-27 20:14:40  
# Original task file: example.txt  
# Log file: ~/greasy/greasy-2010379.log  
#  
# Warning: Task 2 failed  
→ /bin/hostname  
# Warning: Task 13 was cancelled due to a dependency failure  
[# 8 #] /bin/sleep 13  
# Warning: Task 15 failed  
→ /bin/hostname  
  
# End of restart file
```

FICHERO DE REINICIO (II)

- Cambiamos el fichero de tareas original por el de reinicio y volvemos a lanzar el job

```
Start greasing ~/greasy/example.txt.rst
INFO: MPI engine is ready to run with 3 workers
INFO: Allocating task 8
INFO: Allocating task 12
INFO: Task 8 completed successfully on node s01c3b03. Elapsed: 00:00:00
INFO: Allocating task 10
INFO: Task 12 completed successfully on node s01c3b03. Elapsed: 00:00:00
INFO: Task 10 completed successfully on node s01c3b03. Elapsed: 00:00:13
INFO: MPI engine finished
INFO: Summary of 3 tasks: 3 OK, 0 FAILED, 0 CANCELLED, 0 INVALID.
INFO: Total time: 00:00:13
INFO: Resource Utilization: 33.33%
Finished greasing ~/greasy/example.txt.rst
```

TESTING Y SOPORTE

- Solicitar acceso a LaPalma:
cat.cluster.lapalma@iac.es
- Grupo de soporte de LaPalma en el IAC:
res_support@iac.es
- Soporte software GREASY:
support@bsc.es