



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# Planificación adaptativa en Grid: Códigos Monte Carlo

Rafael Mayo García  
CIEMAT

I Jornadas de Computación Distribuida Inteligente  
y Sistemas Complejos

Trujillo, 11 de junio de 2012



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

El trabajo realizado para poder hacer esta presentación supone el eje central de la tesis doctoral de Manuel A. Rodríguez Pascual “Ejecución eficiente de códigos Monte Carlo en infraestructuras Grid” (Madrid, 2012)





GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# Planificación adaptativa en Grid: Códigos Monte Carlo

Manuel Rodríguez Pascual, Rafael Mayo García  
CIEMAT

I Jornadas de Computación Distribuida Inteligente  
y Sistemas Complejos

Trujillo, 11 de junio de 2012



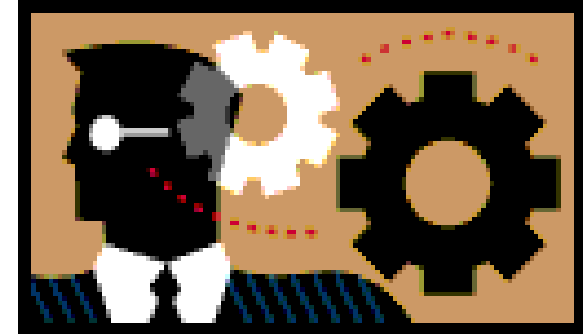
GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

- Problema
  - Justificación
  - Códigos Monte Carlo
- Solución (Montera)
- Resultados
- Conclusiones





- Los códigos Monte Carlo se usan habitualmente en
  - Radiofísica
  - Economía
  - Finanzas
  - Medioambiente
  - Física de Altas Energías
  - Ingeniería
  - Estadística
  - Física de Plasmas
  - Química
  - Técnicas de optimización
  - etc.



- La planificación de tareas en Grid...
  - Ha sido y es objeto de un amplio estudio
  - Técnicas en
    - Replicación de tareas (*M/N, bag of tasks*)
    - Agrupamiento de tareas (*chunk  $\equiv n^{\circ}$  de samples*)
- Pero en general...
  - No se ha aplicado específicamente a códigos Monte Carlo
  - No se ha tenido en cuenta el carácter dinámico de la Grid
  - No se ha estimado el *overhead*
  - Las soluciones se han probado en simuladores



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

Y todo es dinámico en el mundo real...



GOBIERNO  
DE ESPAÑA

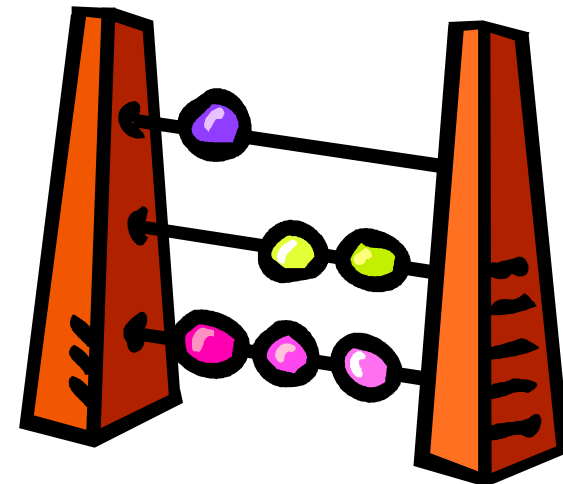
MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# MC: características

- Modelado de problemas complejos
- (Muchas) Simulaciones + Matemáticas
  - números aleatorios para inicializar simulaciones
  - tareas simples e independientes
  - estadísticas para unir los resultados







- Inicialización (I): tiempo constante y/o variable
- Simulación (S): tiempo variable
- Análisis de los resultados (R): tiempo constante y/o variable
- $T_{ejec} = I_c + I_v \cdot N + S_v \cdot N + R_c + R_v \cdot N \sim a \cdot N + b$

(con N = número de *samples*)



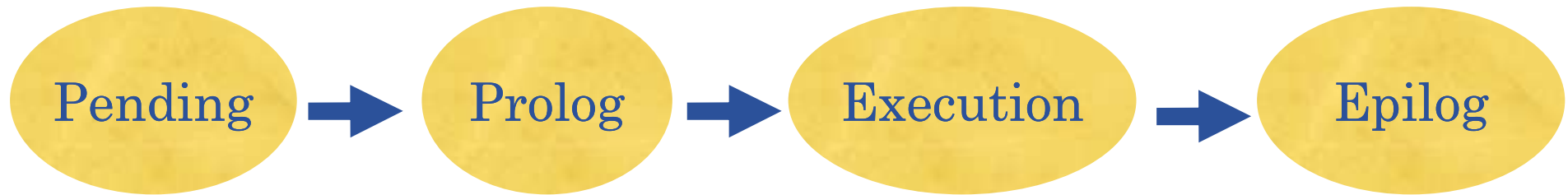
GOBIERNO  
DE ESPAÑA

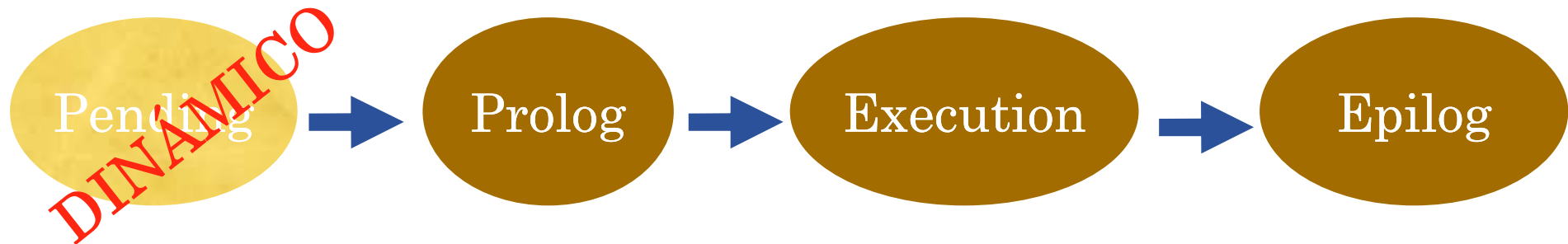
MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

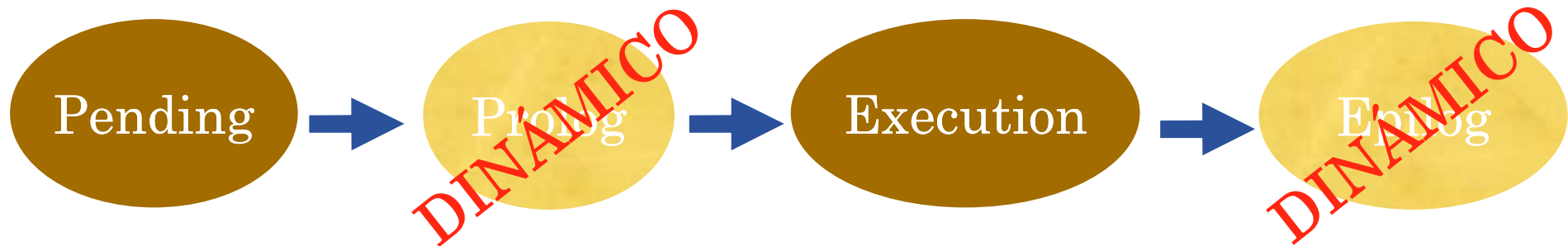
Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# MC: Modelo

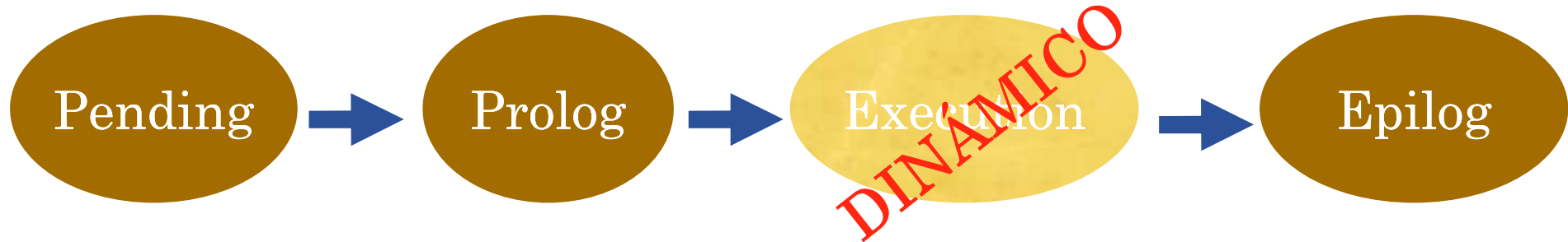




- Status de la infraestructura
  - Dinámico
- Número de tareas a ejecutar
  - Dinámico



- Tamaño de los ficheros de entrada/salida
  - Puede conocerse antes de la ejecución
- Conexión de red local y remota
  - Aproximadamente constante para cada *site*



- Rendimiento del sitio remoto
  - Aproximadamente constante para cada *site*
  - Fluctuaciones dependiendo del uso

- Del código

- Realizar 1, 10, 100, 1000... simulaciones (u otra serie)
- Obtener  $a$  y  $b$ 
  - $T_{ejec} \sim a \cdot N + b$
- Ejecutar el *benchmark* Whetstone en el mismo *site*
- Normalizar
- Repetir este proceso en varios *sites* para evitar distorsiones



- *Del site*
  - Cuándo
    - Cuando se descubre un sitio nuevo o cambia algo de uno conocido (memoria, CPU...)
  - Cómo
    - Whetstone, porque queremos su rendimiento con punto flotante
  - Por Qué
    - Los códigos también están normalizados, así que la estimación del tiempo de ejecución es sencilla y precisa



- *Del site*
  - Globus MDS
  - Después de cada ejecución se calcula de nuevo
    - Rendimiento
      - Eficiencia, tareas fallidas, *slots* disponibles
    - Tiempo de cola
    - Ancho de banda
  - Ponderado
    - Fórmula de Biessel
    - Las ejecuciones nuevas cuentan más que las viejas





GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

Montera

# MONTERA

## (Monte Carlo Rápido)





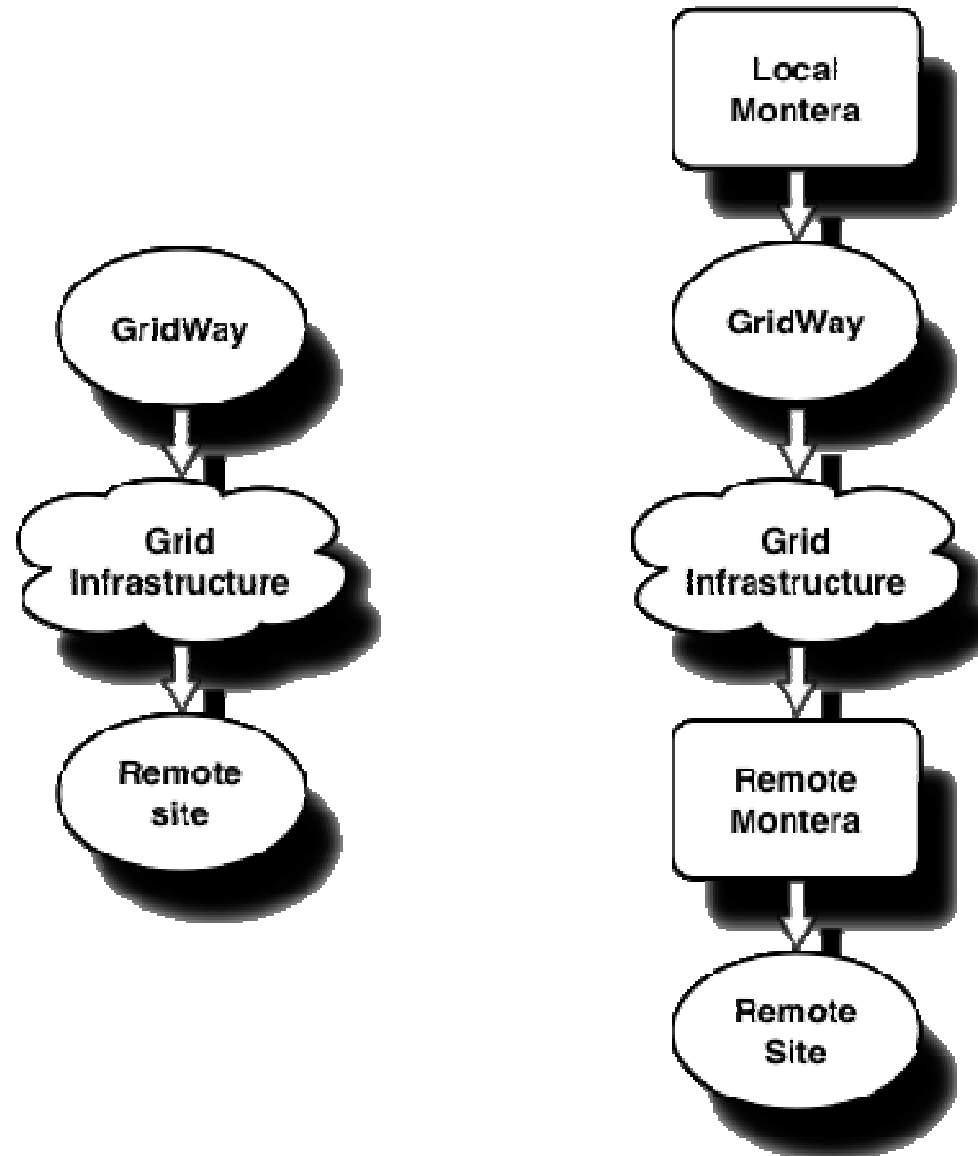
- Software para *scheduling* de códigos Monte Carlo
- Utiliza información estática y dinámica
- Actúa en 2 niveles: recurso local y remoto
- Utiliza
  - GridWay<sup>1</sup>
  - DRMAA<sup>2</sup>

<sup>1</sup> E. Huedo *et al.* Scalable Comp: Prac & Exp **6**, 1 (2005)

<sup>2</sup> P. Tröger *et al.* Proc. CCGrid 2007, 619 (2007)



- **Códigos Monte Carlo**
  - Revisa si son conocidos
  - Realiza un *profiling* si son nuevos
- ***Sites***
  - Revisa si son conocidos
  - Realiza un *profiling* si son nuevos
  - Actualiza la información después de cada ejecución
- **Controla todos los pasos relacionados con la ejecución en Grid del número de *samples* deseado**
  - Puede adaptar diferentes políticas de *scheduling*





- **Montera es una aplicación Java residente en el área de usuario**
  - API para la creación de *chunks* y controlar la ejecución local

- **Plantilla**

```
name=isdep.sh
```

```
num_samples=5000
```

```
input_files=isdep.sh, lgv_LHD_ion_2species.  
files.tar.gz, crea_input, config_mag_LHD.ini
```

```
output_files=output${TASK_ID}.tar.gz
```

```
scheduling_policy=dytss
```

```
max_profiling_time (optional)
```

```
environment_variables (optional)
```



- **Una llamada típica a Montera**

- `Export MONTERA LOCATION=/path/to/Montera`
- `java -Djava.library.path=/usr/local/gw/lib -cp /usr/local/gw/lib/drmaa.jar:/usr/local/gw/include/: /home/manuel/Montera gw03/bin Montera template.mt`

- **Comando final**

- `run.montera template.mt`



- *Dynamic Trapezoidal Self Scheduling*
  - Algoritmo implementado para Montera
    - *Asymptotic performance*
    - *Half performance length*
- Necesita
  - *Profiling* de los Monte Carlo
  - *Profiling* de los Recursos
- **Funcionamiento (simplificado)**
  - Más *samples* a los recursos más potentes (*chunk*)
  - Replicación de tareas para evitar cuellos de botella en tiempo real, no sólo al principio



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# Resultados





- Entorno en producción
  - GridWay 5.4.0
  - Java Virtual Machine 1.5.0.09
  - scheduling interval = 30 s
  - dispatch chunk = 15 jobs
  - maximum number of simultaneous jobs per user =  
= 100



- Códigos Monte Carlo empleados
  - FAFNER2 (simulaciones del orden de segundos) <sup>3</sup>
  - ISDEP (simulaciones del orden de una hora) <sup>4</sup>
  - FastDEP (workflow de los dos anteriores) <sup>5</sup>
  - BEAMnrc (simulaciones del orden de milisegundos) <sup>6</sup>

<sup>3</sup> M. Rodríguez-Pascual *et al.* IEEE TPS **38**, 2102 (2010)

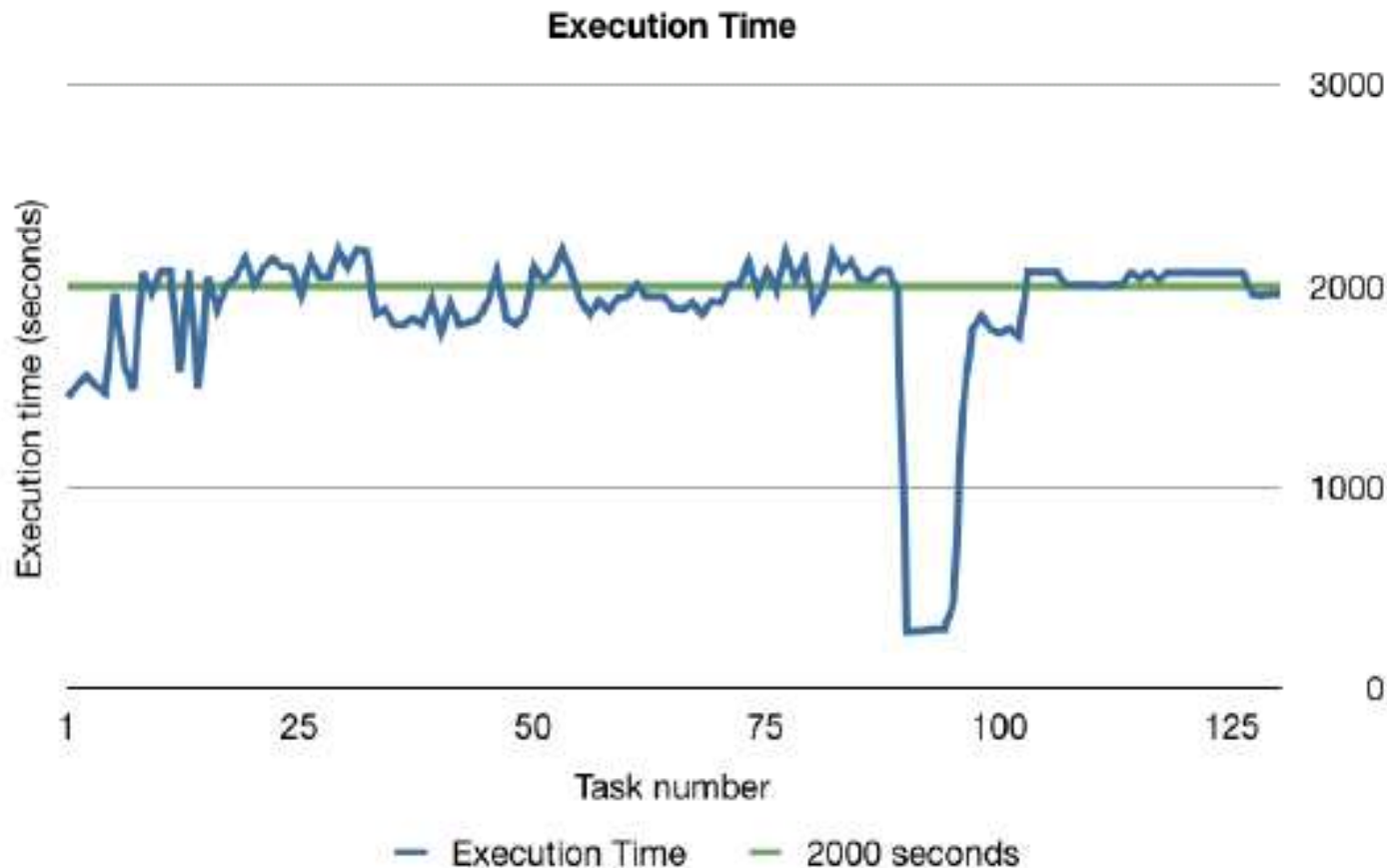
<sup>4</sup> F. Castejón *et al.* Plasma Physics Controlled Fusion **49**, 753 (2007)

<sup>5</sup> M. Rodríguez-Pascual *et al.* CPC (In print)

<sup>6</sup> P. Downes *et al.* Phil. Tran. Royal Soc. A **367**, 2607 (2009)

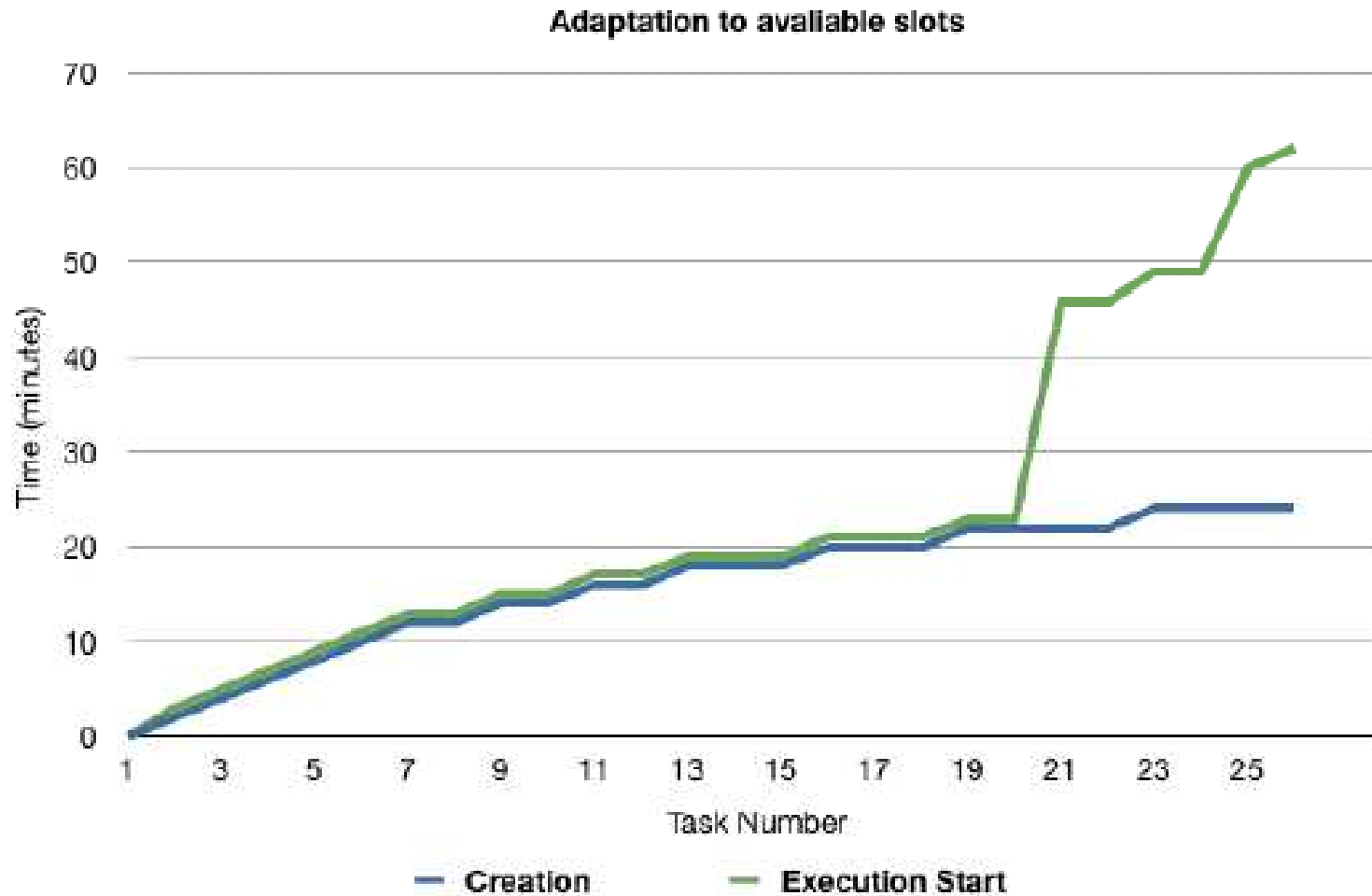


- *Deadline Policy* (entorno real de Producción EGI)
  - Ejecución de tantas tareas en FAFNER2 como sea posible en 2000 s



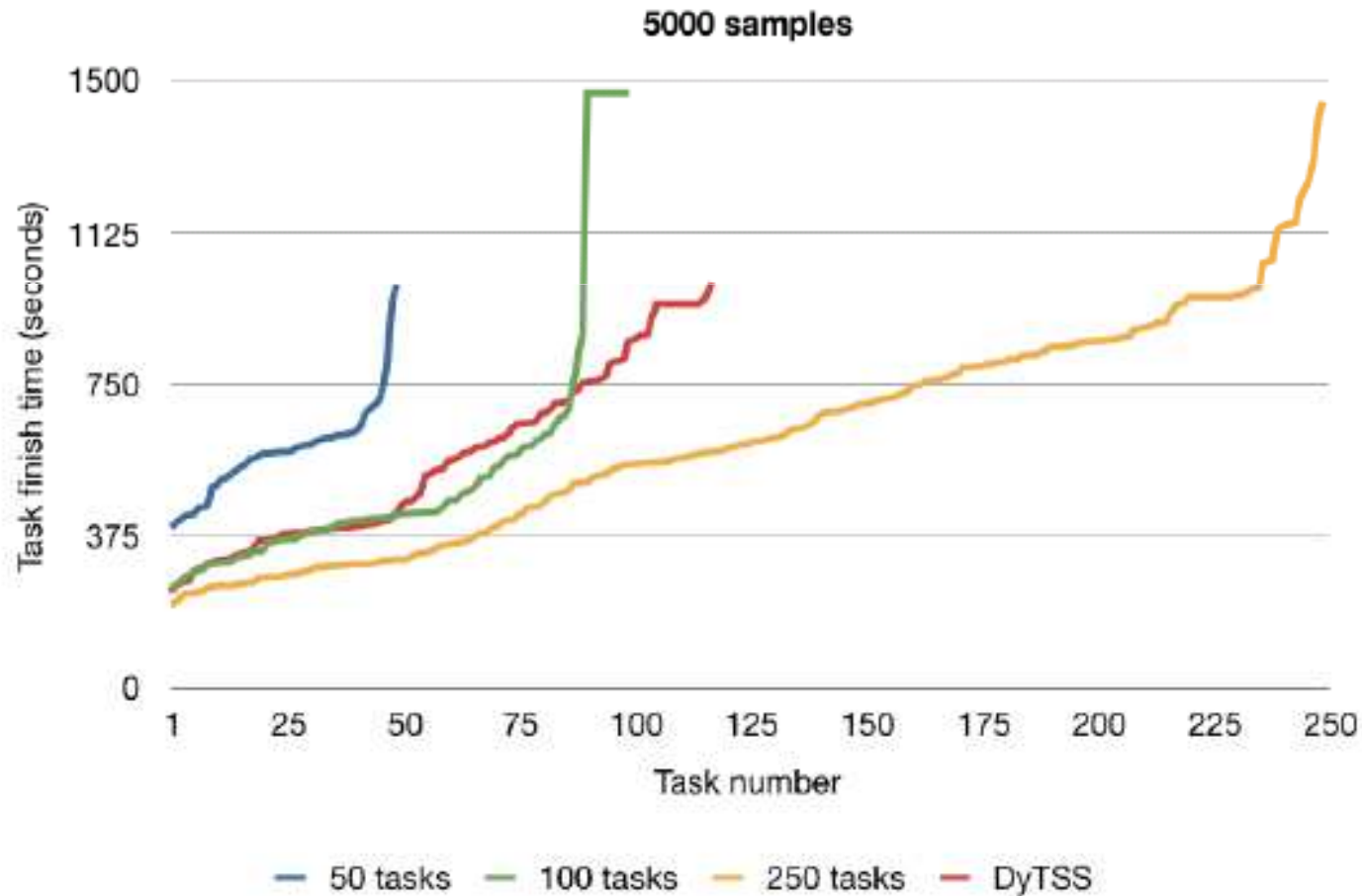


- Adaptación a *slots* libres (`ce01-tic.ciemat.es`)





- Simulador desarrollado para Montera
  - FAFNER2 - 5000 partículas





- DyTSS (FAFNER2 - entorno real de Producción EGI)

<i>Samples</i>	<i>50 tasks (seconds)</i>	<i>100 tasks (seconds)</i>	<i>250 tasks (seconds)</i>	<i>GTSS (seconds)</i>	<i>DyTSS (seconds)</i>
5,000	1331 ± 85,56%	1307 ± 55,77%	5649 ± 65,96%	6892 ± 49,11%	1059 ± 48,91%
25,000	5553 ± 58,79%	5304 ± 62,06%	5568 ± 53,77%	8038 ± 50,03%	2301 ± 43,34%
50,000	5857 ± 61,37%	6131 ± 44,06%	6496 ± 9,88%	7681 ± 15,45%	3385 ± 23,82%
200,000	15665 ± 12,12%	6587 ± 9,07%	10701 ± 29,42%	10456 ± 5,56%	5200 ± 11,57%
400,000	16647 ± 15,17%	14963 ± 25,03%	11568 ± 10,63%	27284 ± 21,98%	11192 ± 12,03%



- **DyTSS (ISDEP - entorno real de Producción EGI)**
  - 5000 partículas (~ 10% de un caso real de estudio)

<i>parameter</i>	<i>250 tasks</i>	<i>500 tasks</i>	<i>GTSS</i>	<i>DyTSS</i>
Execution time [s]	253812	222061	249331	183956
Makespan normalized to DyTSS	1,35	1,21	1,33	1
Standard Deviation [%]	7,04	32,30	29,68	6,96



- **DyTSS (FastDEP - entorno real de Producción EGI)**
  - 5000 partículas (~ 10% de un caso real de estudio)

	Montera	Scheduler/Montera	
	[hh:mm]	GridWay	WMS
makespan	60:26	1,16	1,62





- DyTSS (BEAMnrc - entorno real de Producción EGI)

Simulated samples	Walltime GW / Walltime Montero		
	50 tasks	100 tasks	250 tasks
$10^4$	0,7	0,9	4,6
$10^6$	1,1	1,4	2,4
$10^8$	1,8	1,5	1,3



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

## *Overhead*



- *Overhead* inherente a GridWay
- 300 s para la tarea más larga en el *profiling* del *site*
  - Una única vez
- $10^4$  s para el *profiling* del código
  - Una única vez



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

Conclusiones

# Conclusiones



- Hasta el momento, la planificación de tareas no tenía en cuenta el carácter dinámico de la Grid
- Una planificación adaptativa mejora el rendimiento de los códigos Monte Carlo
- Montera ha sido probado eficientemente en dos códigos Monte Carlo de distinta naturaleza
- Montera ofrece otras posibilidades que no sólo la mera ejecución de tareas
- Montera permite una ejecución desatendida



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

# Conclusiones





GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN

**Ciemat**

Centro de Investigaciones  
Energéticas, Medioambientales  
y Tecnológicas

**MUCHAS GRACIAS**

<http://www.ciemat.es/portal.do?IDR=343&TR=C>