

- CABHALA -

Checkpointing & Limitless en ACME

A. Bustos, A.J. Rubio-Montero, A. Alberto, R. Mayo-García [CIEMAT]

A. Cascajo, D.E. Singh, J. Carretero [UC3M]



Universidad
Carlos III de Madrid



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA
E INNOVACIÓN

Ciemat
Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas

Objetivos

- Monitorizar el consumo energético y el tiempo que cuesta migrar una tarea en HPC
- Optimización de jobs en ejecución mediante migración a nodos más modernos.
- Tareas de administración y mantenimiento.
- Herramientas: LIMITLESS & smigrate
- Infraestructura: clúster de desarrollo ACME

LIMITLESS

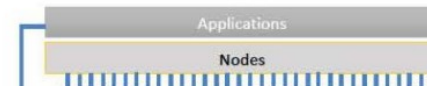
LIMITLESS - Light-weight MonItoring Tool for Large Scale Systems

Alberto Cascajo
University Carlos III of Madrid
Leganés, Madrid, Spain
acascajo@inf.uc3m.es

David E. Singh
University Carlos III of Madrid
Leganés, Madrid, Spain
dexposit@inf.uc3m.es

Jesus Carretero
University Carlos III of Madrid
Leganés, Madrid, Spain
jcarrete@inf.uc3m.es

Abstract—This work presents LIMITLESS, a HPC framework that provides new strategies for monitoring clusters. LIMITLESS is a scalable light-weight monitor that is integrated with other HPC runtimes in order to obtain an holistic view of the system that combines both platform and application monitoring. This



- Monitor HPC – mide energía consumida por cada procesador en el cluster, entre otras magnitudes (A. Cascajo *et al*, UC3M)
- Servidor central y demonios ligeros en los nodos de cálculo. Medida cada 5s
- Memory leak a las ~24h después de activarlo (bug reportado) 🍌

May 19 02:38:27 acme kernel: Out of memory: Kill process 116934 (Test_server_IB_) score 937 or sacrifice child

May 19 02:38:27 acme kernel: Killed process 116934 (Test_server_IB_), UID 5316, total-vm:102565840kB, anon-rss:90745824kB, file-rss:0kB, shmем-rss:0kB

checkpointing - smigrate

- Integra DMTCP y Slurm para migrar tareas entre nodos de un mismo cluster (M. Rodríguez-Pascual *et al*, CIEMAT).
- Independiente de la aplicación
- DMTCP da problemas cuando el ejecutable necesita abrir y cerrar archivos para volcar el output. Restringe un poco las posibilidades.
- DMTCP no funciona entre nodos con **distinto hardware y/o SO**. Mirando las FAQs (<https://dmtcp.sourceforge.io/FAQ.html>):
“... Note that in migrating between arbitrary computers, it is possible to encounter "illegal instruction" on restart. This can occur if the CPUs of the source machine and destination machine are different. ...”
- Esto elimina la posibilidad de migrar tareas para mejorar el rendimiento.

The Journal of Supercomputing
<https://doi.org/10.1007/s11227-021-03663-1>



Performance and energy task migration model for heterogeneous clusters

Esteban Stafford¹ · José Luis Bosque¹

Accepted: 28 January 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

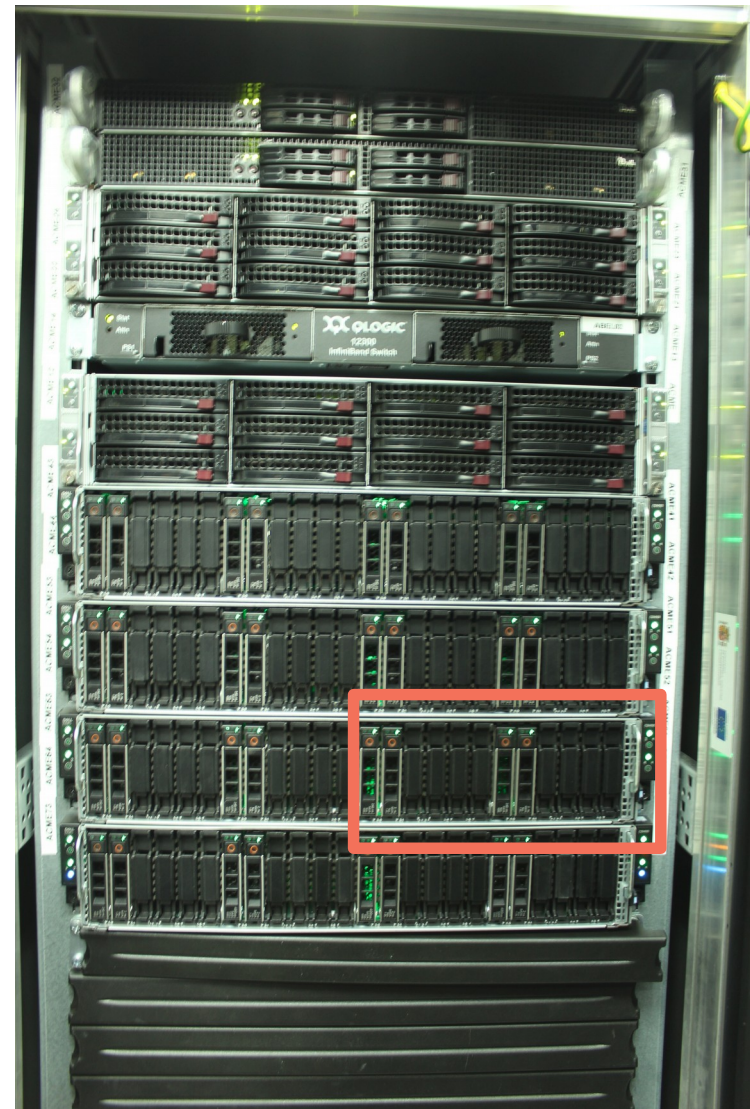
This article presents a set of linear regression models to predict the impact of task

Códigos

- Ejemplos de la suite BOTS del BSC
[<https://pm.bsc.es/gitlab/benchmarks/bots>]
- Varios ejecutables OpenMP
 - nqueens
 - sparselu
 - fft
- Tamaño del problema fácilmente variable
- Dos casos: small y normal

Infraestructura

- ACME – cluster de desarrollo del Dep. Teconología del CIEMAT.
- Monitor LIMITLESS en wndevel, demonios en acme61 y acme62.
- 40p - Intel(R) Xeon(R) Gold 6138 CPU @2.00GHz
- Partición exclusiva para estos experimentos.

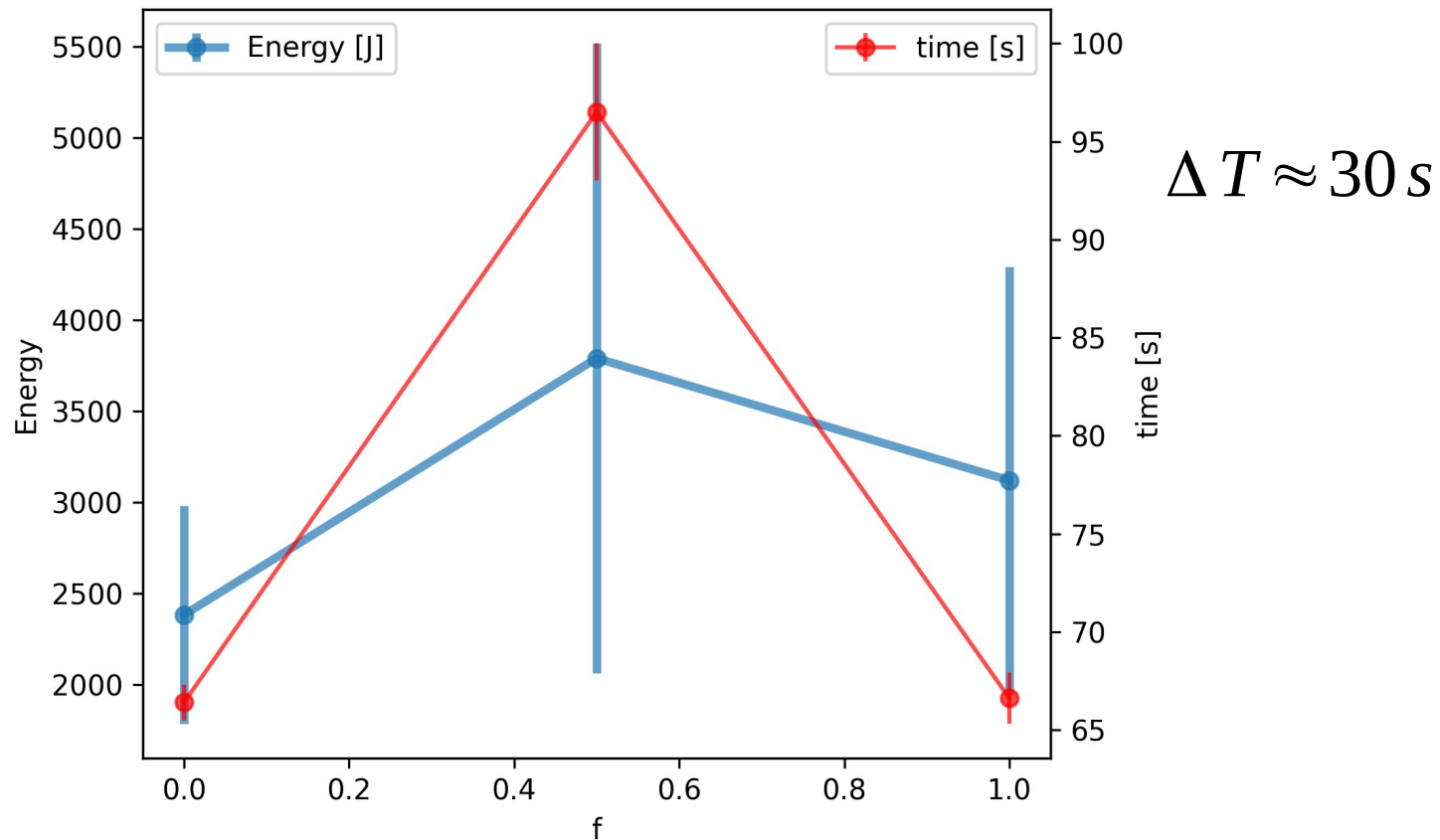


Metodología

- Script en Python para lanzar jobs, migrarlos y guardar información relevante.
- Script en Python para recopilar esta información, cruzarla con el log de Limitless y sacar estadísticas de tiempos de ejecución y energías consumidas.
- f = punto en la ejecución en el que se migra:
 - $f=1$ → no se migra (nodo1),
 - $f=0$ → no se migra (nodo2),
 - $f=0.5$ → migración a la mitad de la ejecución

n-queens (small)

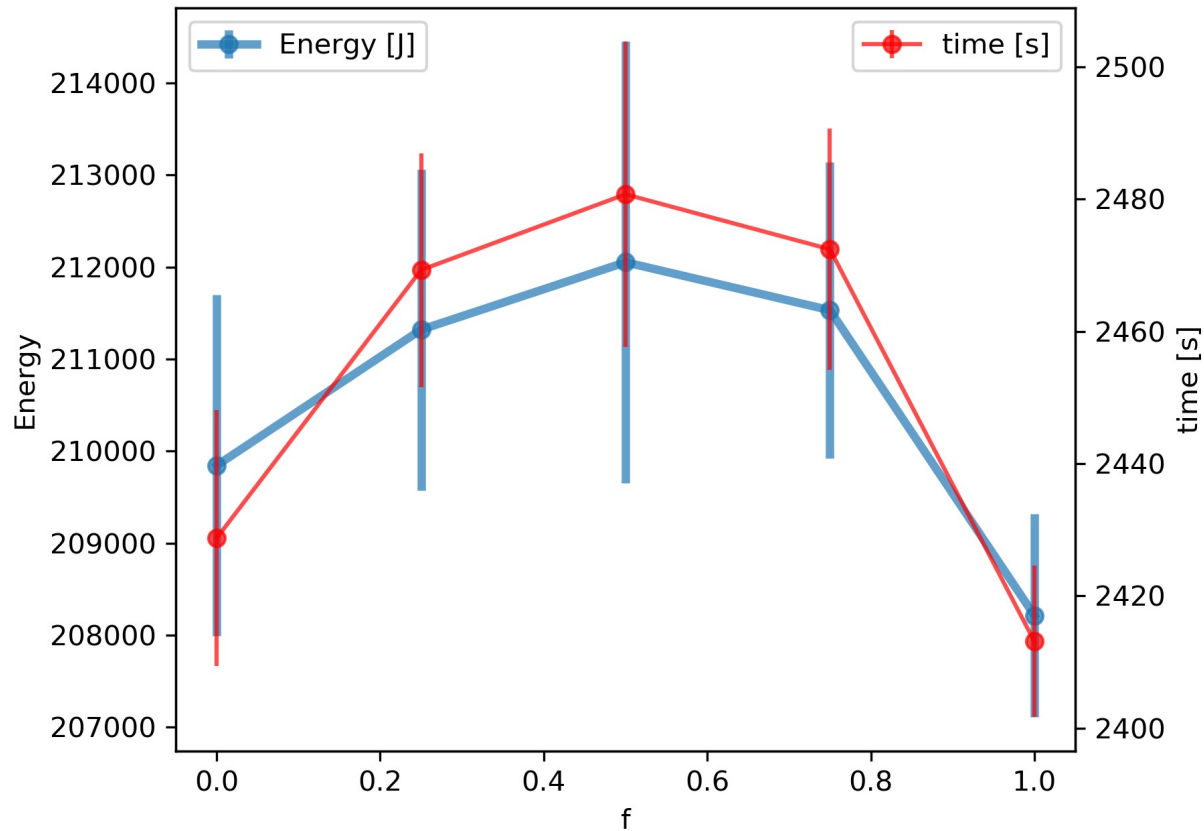
- $N=16$, $np=16$, el error es la std de ~ 30 ejecuciones



- Overhead temporal visible, comportamiento extraño en la energía.

n-queens

- N=18, np=40, el error es la std de 10-12 ejecuciones

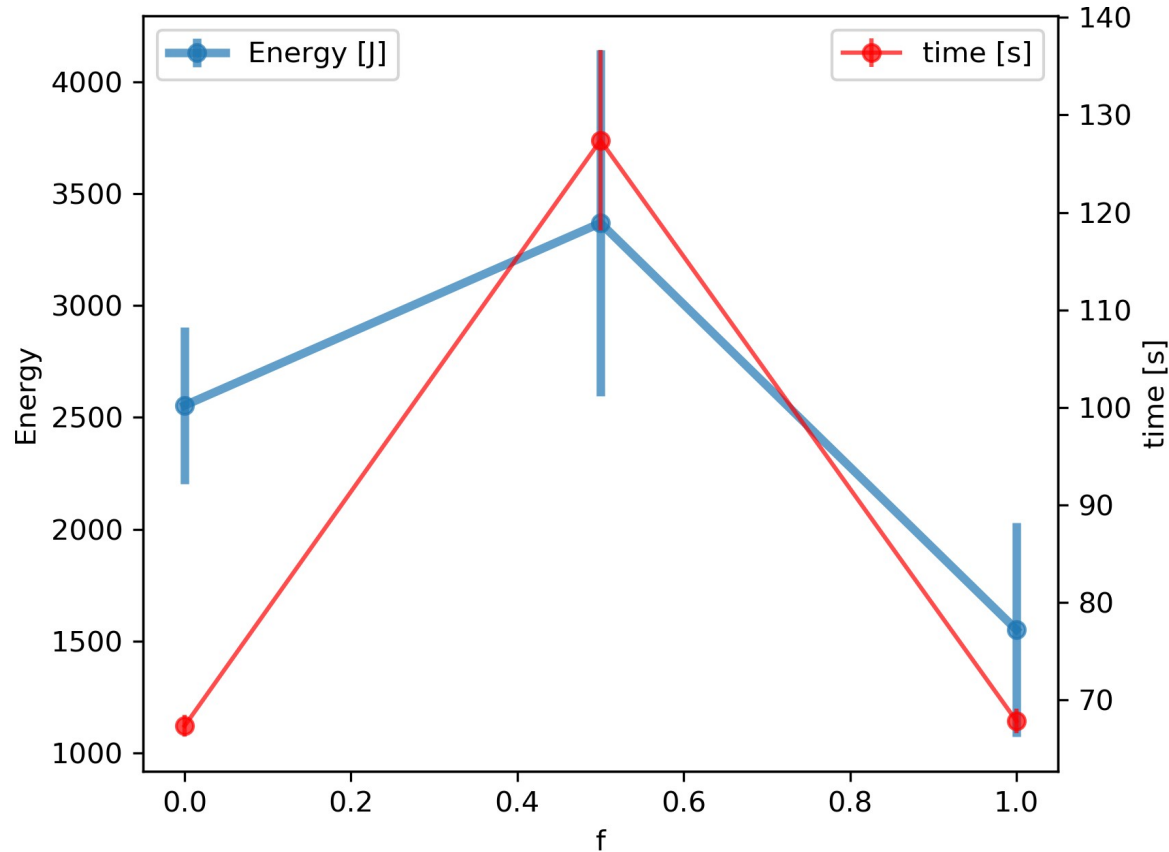


$$\Delta T \approx 70s$$

- Datos muy dispersos, hay tendencia pero poca diferencia.

sparselu - small

- $n=m=150$, $np=16$, el error es la std de ~ 50 ejecuciones

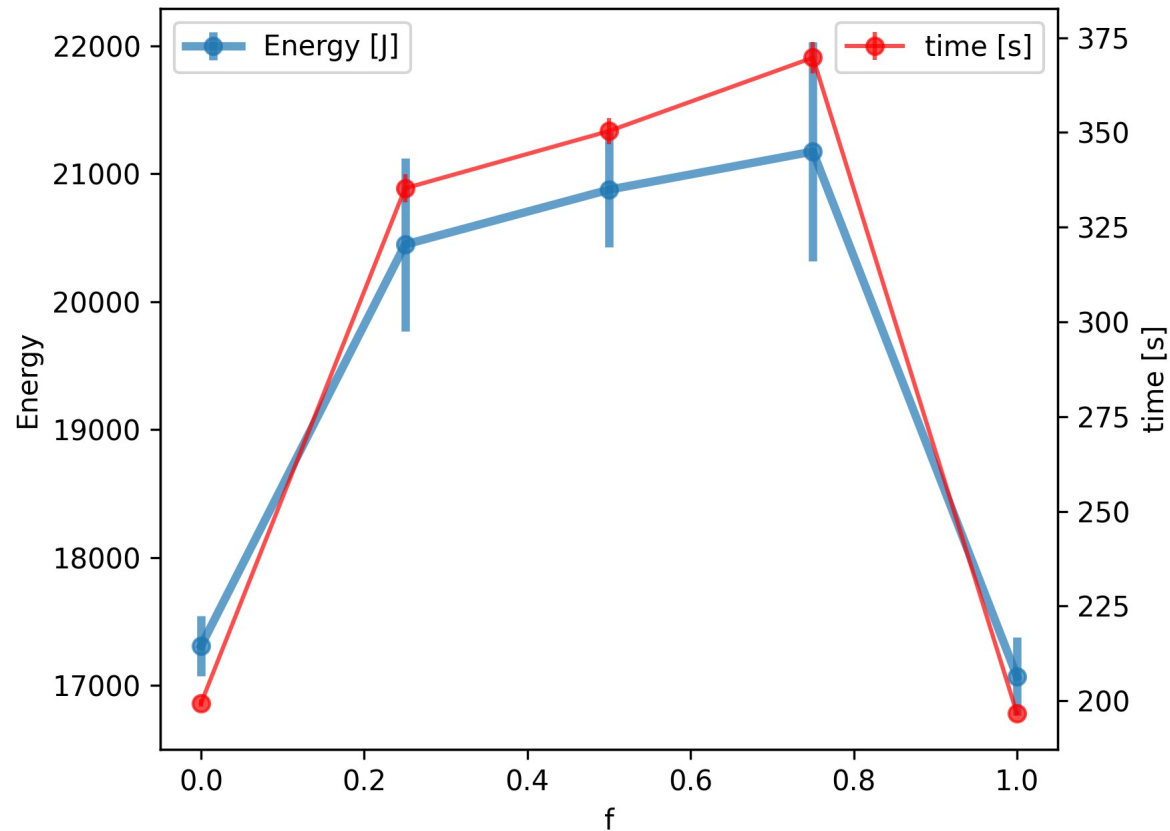


$$\Delta T \approx 60 \text{ s}$$

- Overhead temporal y en la energía visibles.
Comportamiento raro (marginamente) en la energía

sparselu

- $n=m=200$, $np=40$, el error es la std de ~ 20 ejecuciones

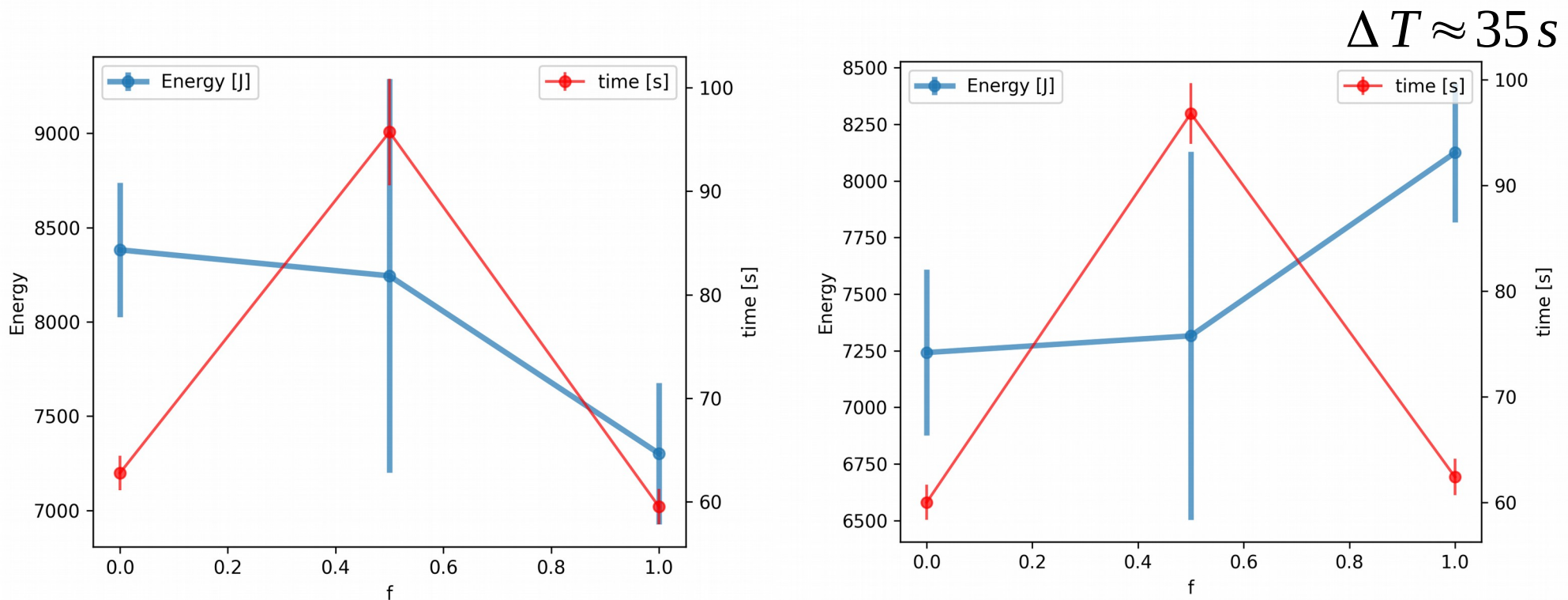


$$\Delta T \approx 150 \text{ s}$$

- Overhead temporal y en la energía visibles, y cierta dependencia con f cuando hay migración.

fft - small

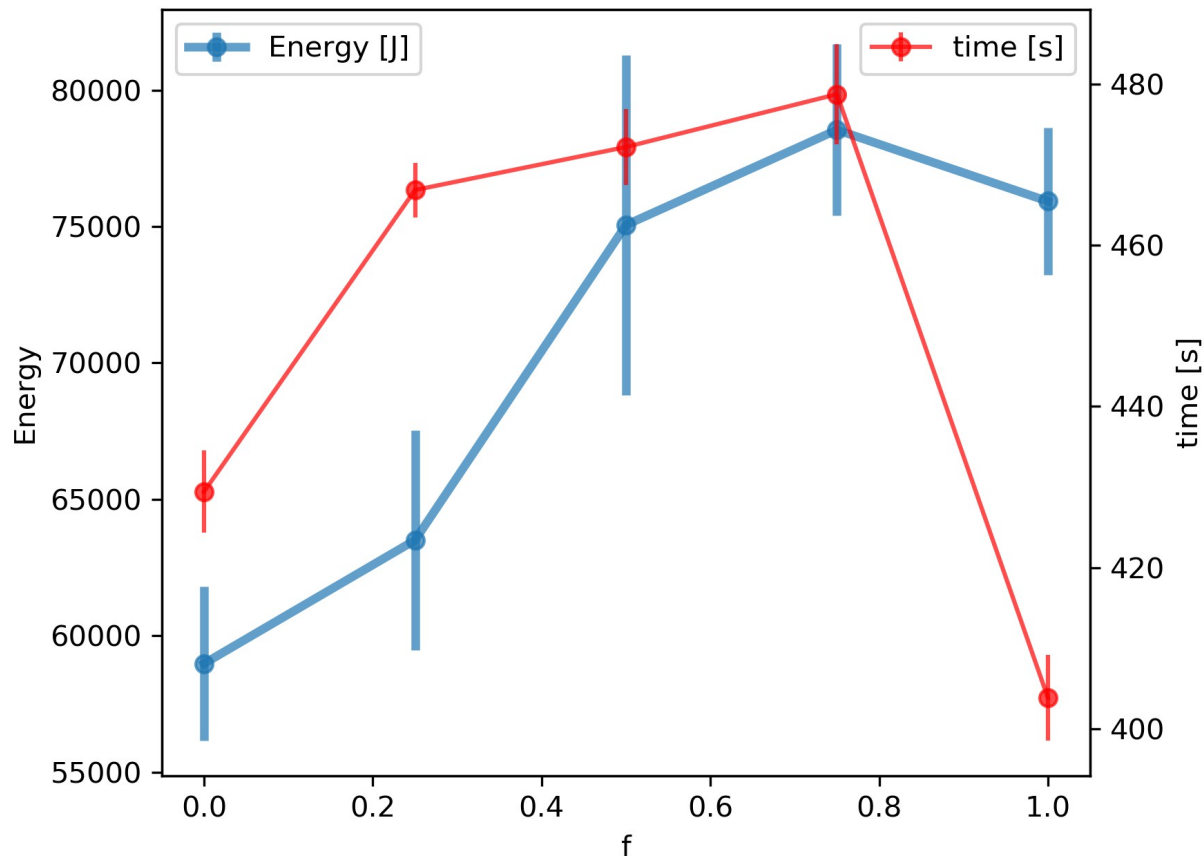
- $n=5000000$, $np=16$, el error es la std de ~ 50 ejecuciones
- Para comprobar, intercambiamos los nodos origen y destino.



- Overhead temporal visible. Comportamiento raro en la energía, dos nodos idénticos consumen ligeramente diferente.

fft

- $n=20000000$, $np=40$, el error es la std de 10 ejecuciones



$$\Delta T \approx 70s$$

- Overhead temporal visible. Comportamiento raro en la energía, dos nodos idénticos consumen ligeramente diferente. Dependencia con f .

Conclusiones

- No se han podido lograr los objetivos iniciales del proyecto.
- Buenas medidas del tiempo de migración, dependiente de la aplicación y de sus parámetros de entrada.
- Para jobs pequeños, mucha dispersión en las medidas, sobre todo de E.
- En conjunto, imposible identificar un comportamiento universal. Dos nodos idénticos pueden consumir distinta E haciendo la misma tarea y también depende del punto de migración [f] – ojo con la optimización de ejecuciones!
- Útil para tareas de administración de HPC, quizás no tanto para usuarios.