Supercomputación y E-Ciencia WP7 Engineering / Turbulence

Guillem Borrell i Nogueras

@guillemborrell

http://guillemborrell.es http://torroja.dmt.upm.es



ETSI Aeronáuticos

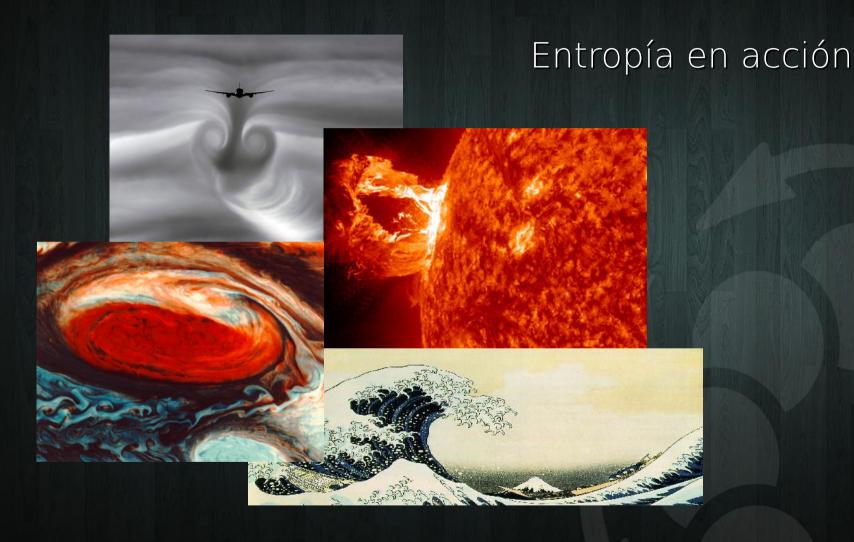


Universidad Politécnica de Madrid

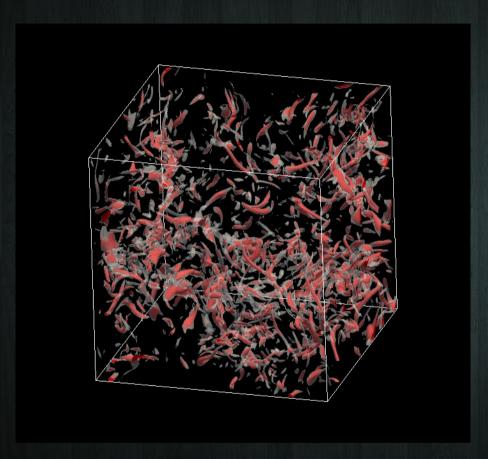
Antes de empezar...

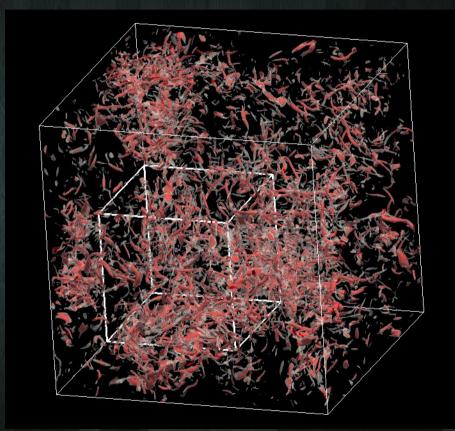
¿Por qué es importante entender la turbulencia y por qué usamos grandes cantidades de recursos para ello?

Donde hay un fluido, habrá turbulencia.



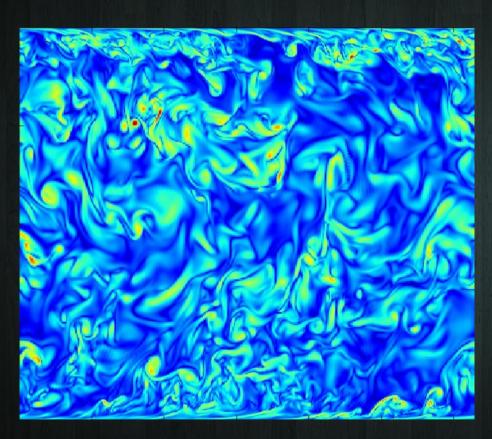
El tamaño importa

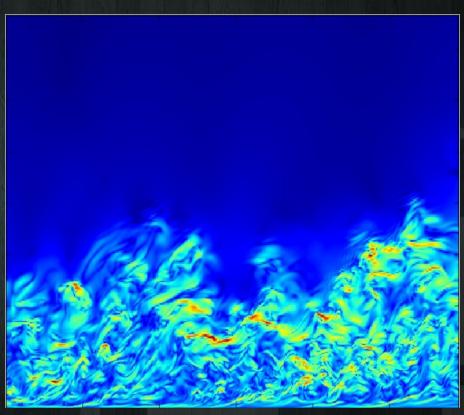




El número de Reynolds es la cantidad de información por unidad de volúmen

"La Turbulencia" no existe.





Aún así tampoco entendemos del todo los casos canónicos.

Tenemos un canal en Youtube!

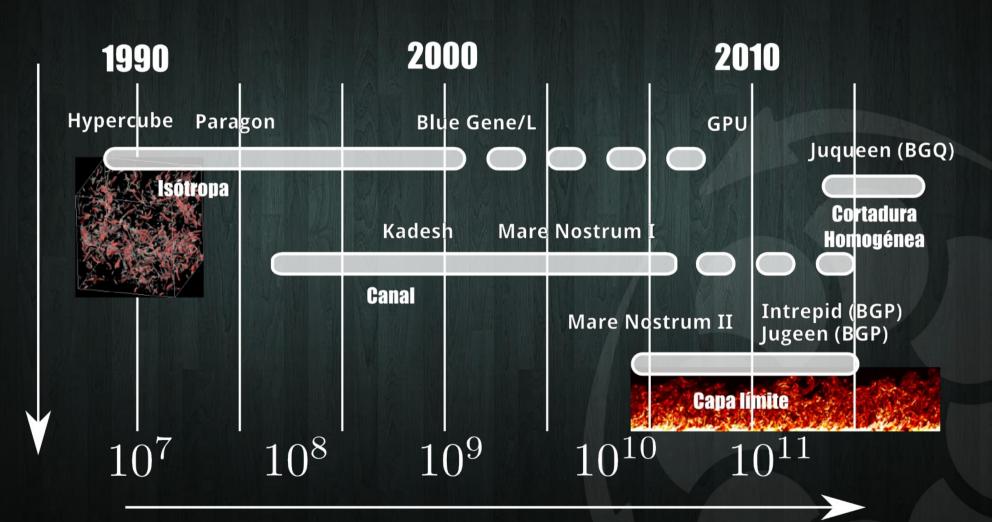
http://www.youtube.com/user/LMFCatUPM

Problemas...

- ¿Cómo conseguir un archivo que tiene 21 años?
- ¿Cómo mover 75 TB desde Jülich a Madrid?
- ¿Cómo visualizar un archivo de 120 GB?

En estos tres casos el Consolider ha participado directa o indirectamente.

Nuestro parámetro es el tiempo, y empieza hace 20 años.



Somos un grupo pequeño, no tenemos gente de IT.

iRods

- Irods.org
- Jaakko Lappalainen (UPM) y Toni Cortés (BSC)
- Base de datos para metadatos.
 - Dónde está el archivo (BSC, TACC, ANL, UPM...)
 - Qué propiedades tiene.
 - Cómo se lee.
 - Quién lo escribió.
- Estos archivos fueron utilizados para artículos ya publicados. Generar archivos nuevos es peor idea de lo que parece.
- No elimina algo de ingeniería inversa, pero hace que ciertas cosas sean posibles.

En 2010/11 generamos los datos.

- INCITE. ANL. 80 MHCPU 32K cores BG/P
 - 220 TB de datos.
- PRACE, JFZ. 40 MHCPU 32K cores BG/P
 - 75 TB de datos.
- Los centros de supercomputación son para correr.
 - Almacenar datos suele ser una actividad adicional.
 - Mover datos suele ser una actividad adicional (y traumática).
- He acabado la simulación. ¿Y ahora qué?

¿Y ahora qué?

- Mover los datos de ANL al TACC.
 - Archivo.
 - Post en el TACC
 - Mando datos a Madrid por BBCP.
- Mover los datos de JFZ al BSC.
 - Archivo.
 - Pero el BSC está en transición MN II III.
 - No hay suficiente ancho de banda en Madrid para los dos proyectos.
- 2010-2012: 60% del tráfico de la UPM.

Comunicación de alta latencia y gran ancho de banda.

- Mandamos discos al BSC.
- Support se encargó de todo.
 - Agradecidos estamos a David Vicente y Javier Bartolomé (BSC).
 - Pusimos el hardware, se lo mandamos y se encargaron de todo.
 - Gestión para realizar el backup y dejar una copia en cinta.
 - Poner hardware externo en un centro de supercomputación es como que se congele el infierno.
- Y los mandaron de vuelta.
- No fue una experiencia agradable.
 - Requiró burocracia y es algo que podría mejorarse significativamente.

Ver es creer.



Paraview funciona muy bien pero...

Los vóxeles están de moda



Mejorar la visualización de grandes escenas.

- Carlos Duelo (UPM/CESViMa)
- SVO (Sparse Voxel Octree).
 - Es multiresolución.
 - Rápido, 0(log N).
 - Muy eficiente en memoria.
 - Lleva ahí 40 años, pero como no sirve para videojuegos...
- Soporte multiGPU en host único.
- Soporte multiGPU en múltiples hosts.
- Soporte CPU-GPU caching.

Y funciona! (Más o menos)

