



# *Informática Científica en el CERN*

*Al servicio de la Ciencia*

*Dr. Miguel Angel Marquina  
Spanish Teachers Programme*

*27 de junio de 2011*



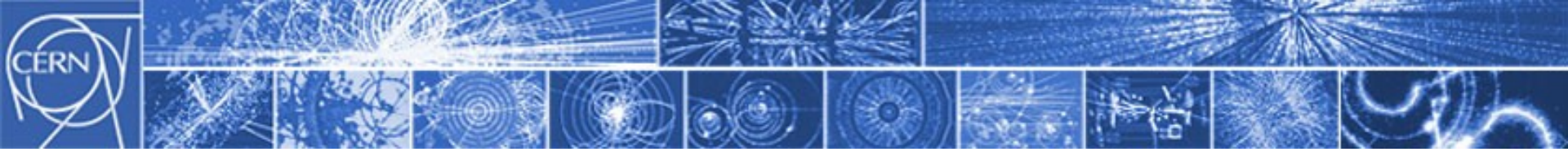
# *Nuestros ejes de actividad*

**Formación**

**Colaboración**

**Investigación &  
Descubrimientos**

**Tecnología**



# Viaje al Mundo de la Información

Cuanto alberga...



# Viaje al centro de la Información -1

- La final de la Sub 21  
1 bit
- Una letra cualquiera  
8 bits= 1 byte
- Una palabra  
10 bytes
- Un SMS  
100-160 bytes

- Una página de texto  
1-2 Kilobytes (Kb)
- Una página de una enciclopedia  
10 Kb
- Una foto de baja resolución  
100 Kb
- El texto de un libro típico  
1000 Kb = 1 Megabyte (Mb)



## *Viaje al centro de la Información -3*

- Un minuto de música (MP3)  
1 Mb
- Un diskette  
1.44 Mb
- Una foto de alta resolución  
2 Mb
- Un minuto de música (CD), radiografía  
10 Mb

- Un diccionario enciclopédico  
100 Mb
- Un CD  
700 Mb
- Un iPod con 200 canciones  
1000 Mb = 1Gb
- Un DVD  
4.7 Gb
- Toda la música de Michael Jackson  
20 Gb



## Viaje al centro de la Información -5

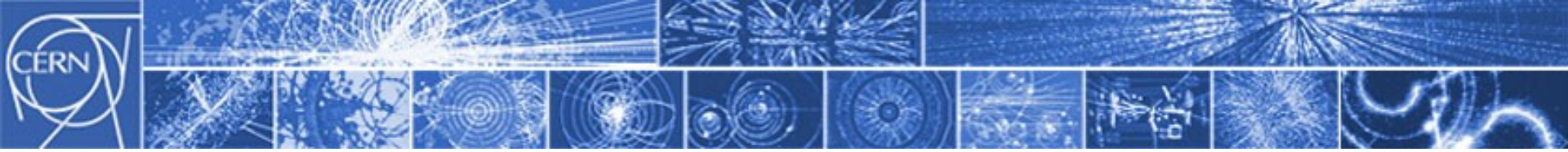
- Una biblioteca especializada  
1000-2000 Gb=1-2 Terabytes
- La Biblioteca Nacional  
8.4 Tb
- La mayor biblioteca del mundo (LOC)  
21 Tb
- Texto producido anualmente en el mundo  
160 Tb
- Lo que alberga Google  
2000-5000 Tb=2-5 Petabytes





## *Viaje al centro de la Información -6*

- Todo el correo electrónico en el mundo  
11 Pb/año
- Las emisiones de TV  
14 Pb/año
- El correo postal  
150 Pb/año
- Las llamadas telefónicas  
600 Pb/año

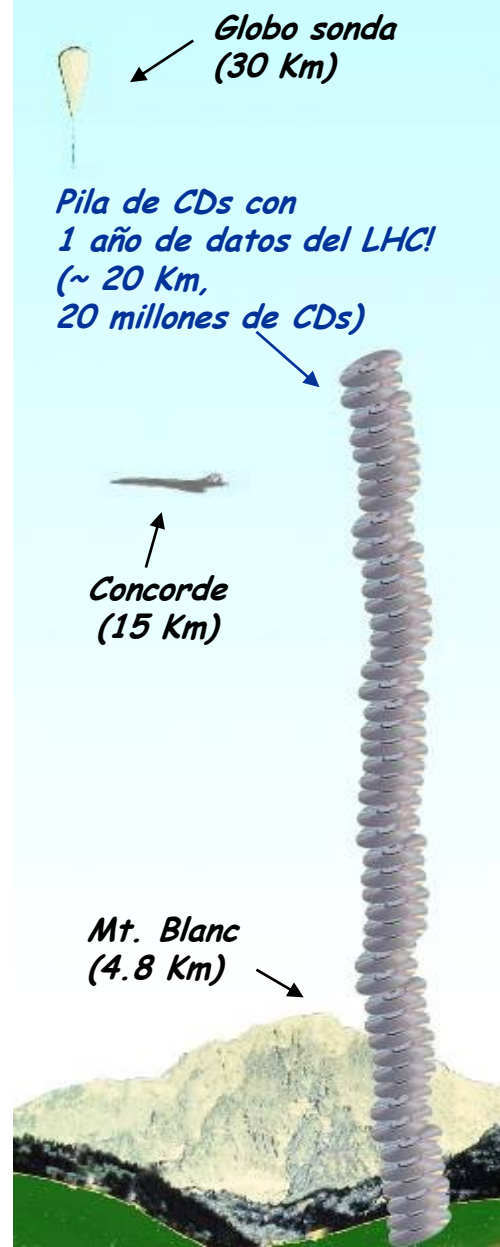


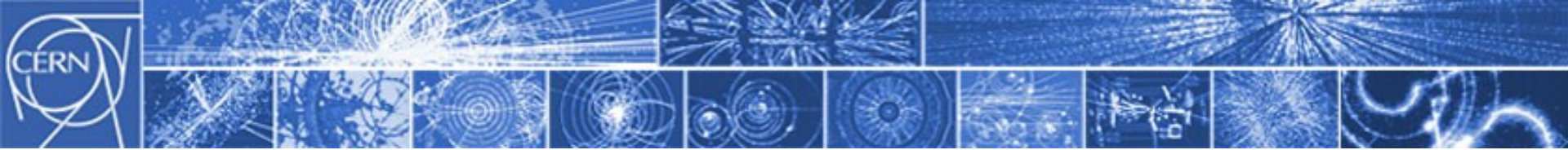
Cuanta información  
produciremos en el  
CERN?



# Los datos del LHC

- 600 millones colisiones/segundo
- 15 Petabytes/año de datos, que son...
  - 1.8 Tb/hora = 30 Gb/min = 500Mb/seg
  - Casi 1 CD cada segundo
  - 6 DVDs cada minuto
  - 1 Biblioteca Nacional cada 5 horas
  - 2 LOC cada día!



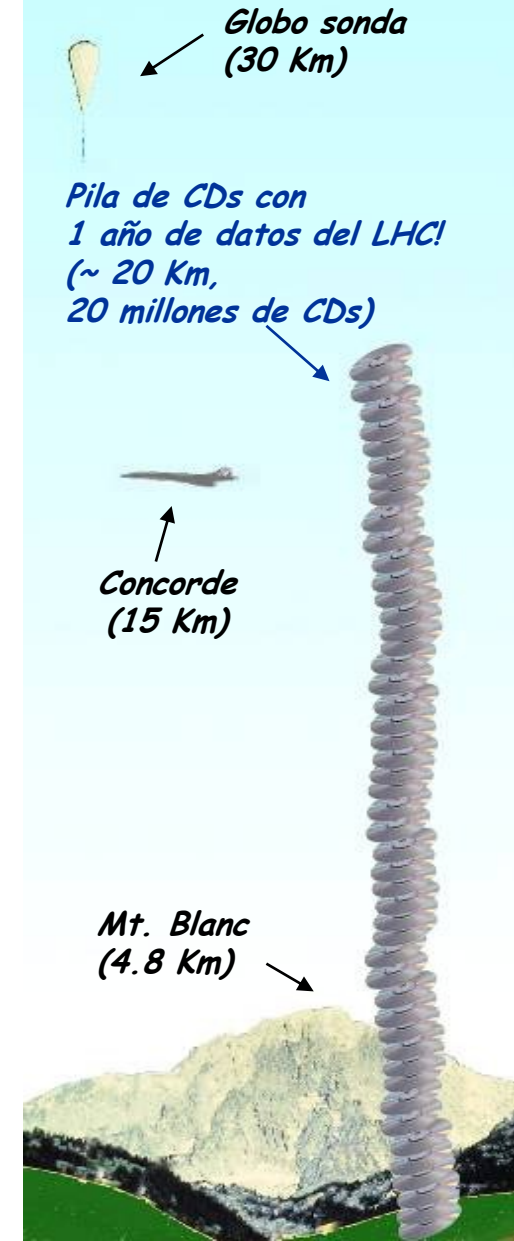


# Los datos del LHC

- El análisis de los datos requerirá un poder de cálculo equivalente a 100000 PCs
- En el centro de Cálculo del CERN tenemos:
  - Más de 16000 PCs
  - Capacidad de almacenamiento de más de 16PB en cinta magnética, 5PB en disco

**¡ Podemos asumir la quinta parte de la capacidad necesaria!**

**De donde sacamos el resto??**



# Compartiendo piso...

- Uno solo no puede asumir los gastos del alquiler de un piso
- Entre varios acuerdan compartirlo
- Se aporta según un acuerdo (espacio/tiempo de habitación/poder adquisitivo...) a un fondo común
- Se consume proporcionalmente a la aportación... o se aporta según consumo

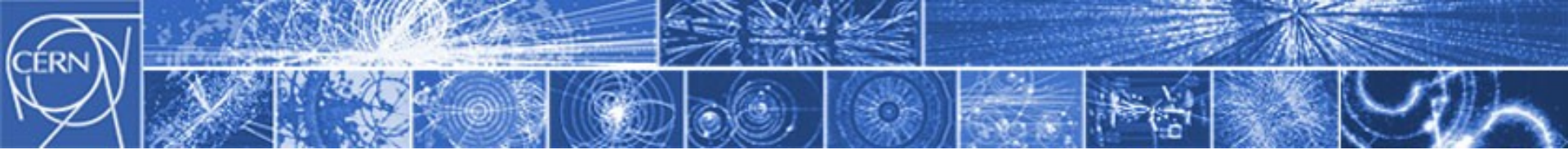
Con las necesidades de computación  
ocurre parecido



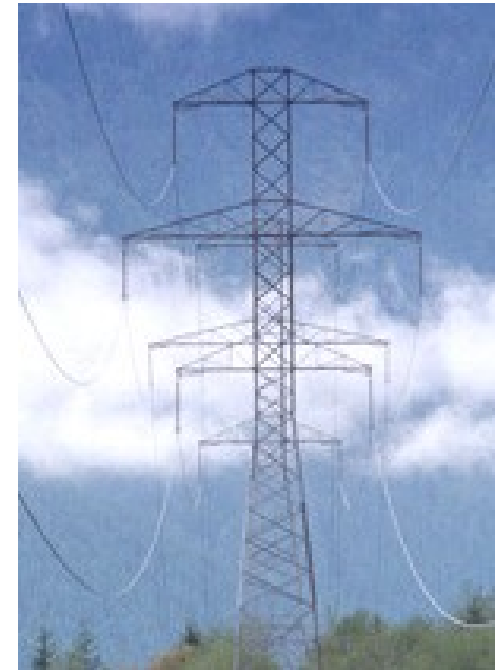
# Calculando a Escote

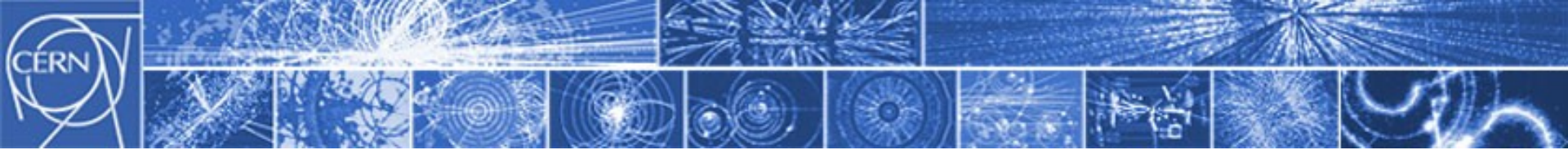
- Uno solo no puede asumir la necesidad de cómputo de un proyecto
- Entre varios acuerdan compartirlo
- Se aporta según un acuerdo (cantidad de recursos disponibles) a un fondo común
- Se consume proporcionalmente a la aportación

Eso es la GRID... en esencia



- El nombre lo acuñaron Ian Foster and Carl Kesselman
- Por analogía a la acepción utilizada en USA para la red eléctrica:  
Electric power grid
- También conceptualmente: no es importante su origen sino la garantía de acceso

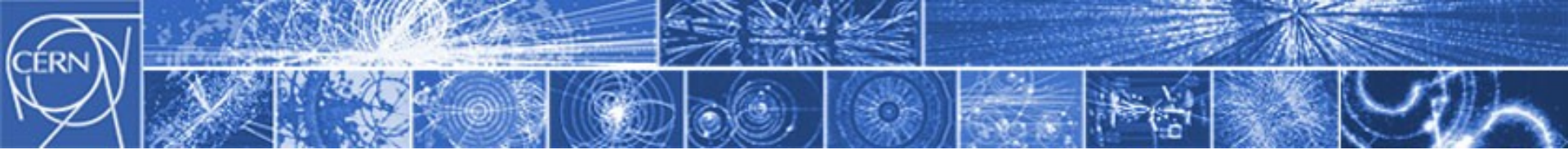




- La Grid utiliza un avanzado software (**middleware**), que asegura la comunicación entre ordenadores en diferentes partes del mundo.
- El Grid **no sólo encontrará los datos** que los científicos necesiten, si no **también las técnicas para procesarlos y la capacidad de cálculo necesaria.**
- Distribuirá la tarea a procesar a **aquella parte en el mundo donde haya capacidad disponible,** y enviará el resultado al científico.







## The GRID middleware:

- Encuentra los lugares convenientes donde se ejecutará el encargo de cálculo enviado por el científico.
- Optimizará el uso de recursos dispersos alrededor del mundo.
- Organizará eficientemente el acceso a los datos científicos.
- Se encarga del proceso de autenticación en los diferentes sitios que el científico usará y de sus políticas de asignación de recursos.
- Ejecutará las tareas.
- Monitorizará su progreso

... y finalmente....

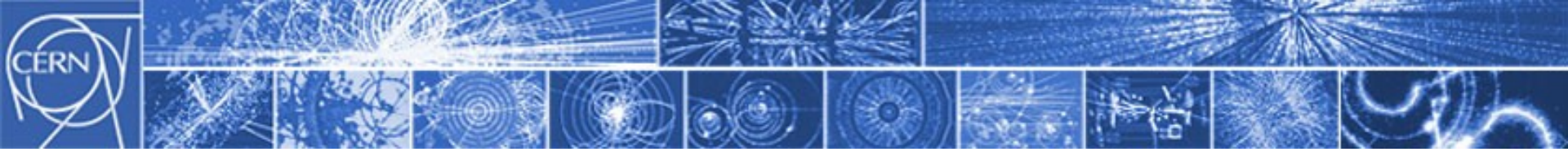


**¡ Informa cuando el trabajo se ha completado y transfiere su resultado !**

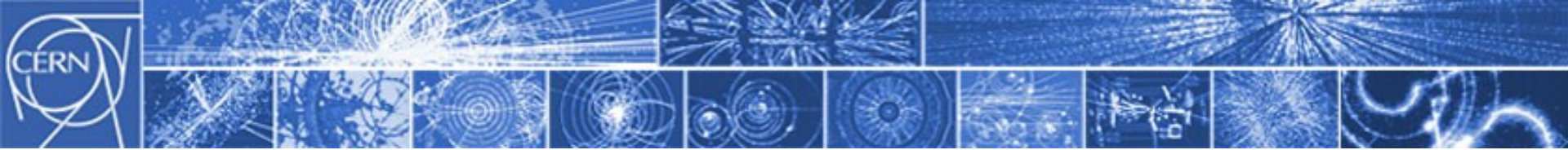
# Web y Grid

- La web facilita a las personas **ACCEDER** a ordenadores dispersos por todo el mundo para **COMPARTIR** información
- La Grid facilita a los científicos **COMPARTIR** (centros con) ordenadores dispersos por todo el mundo para **PROCESAR** información





# ¿Y qué consigue la sociedad con todo esto?



# World Wide Web

Fue desarrollado en  
el CERN, en 1989 !

Sin derechos de  
autor 😊

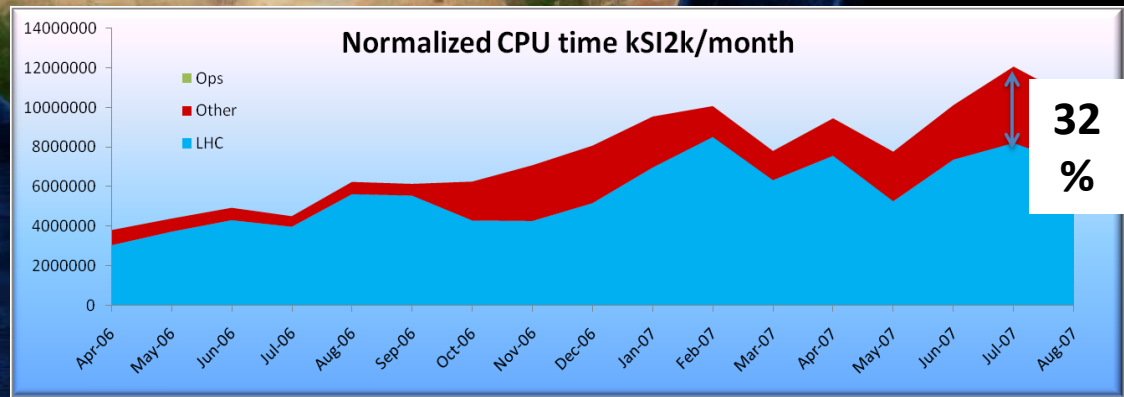
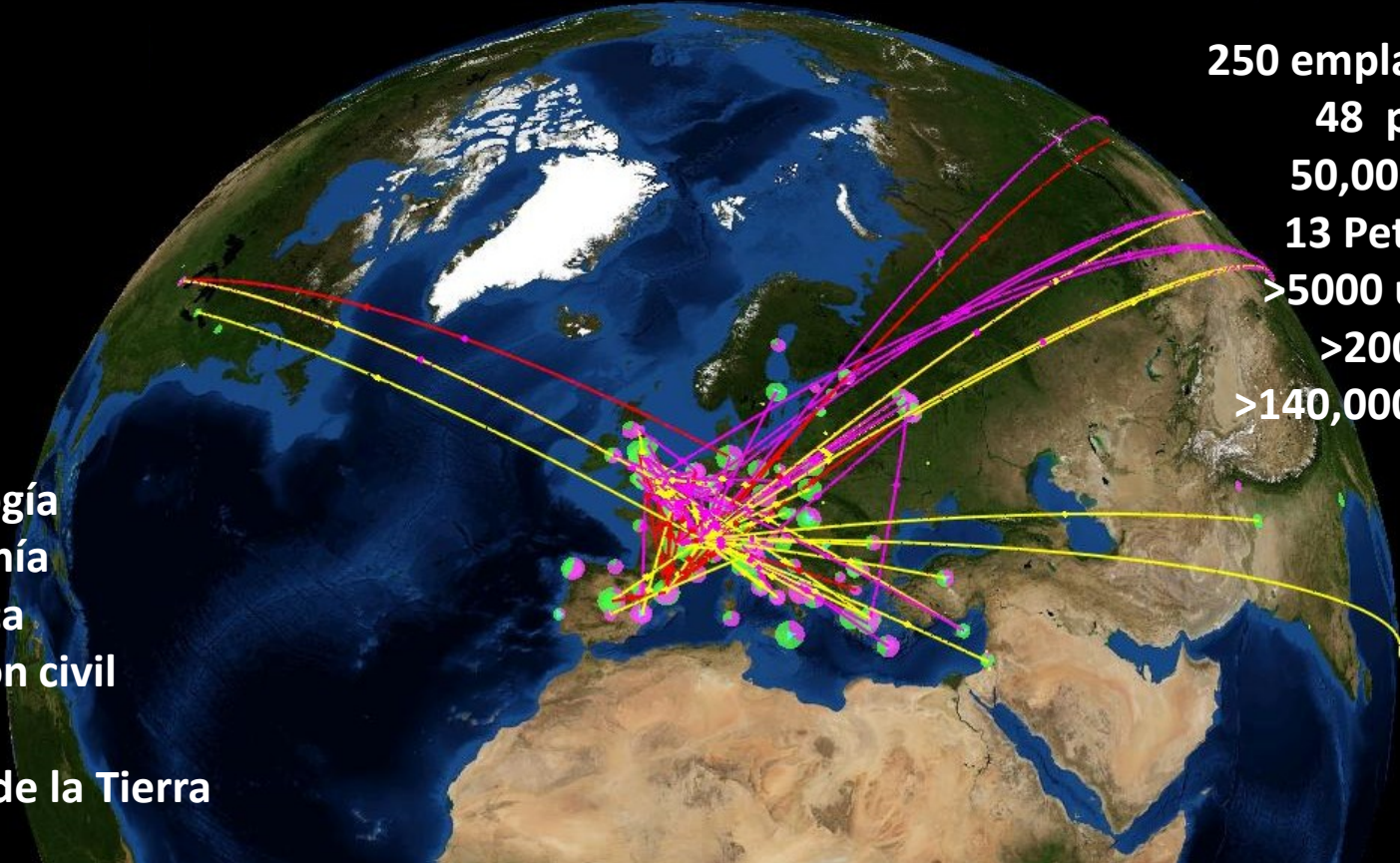




Enabling Grids  
for E-science

**250 emplazamientos**  
**48 países**  
**50,000 CPUs**  
**13 PetaBytes**  
**>5000 usuarios**  
**>200 VO's**  
**>140,000 jobs/día**

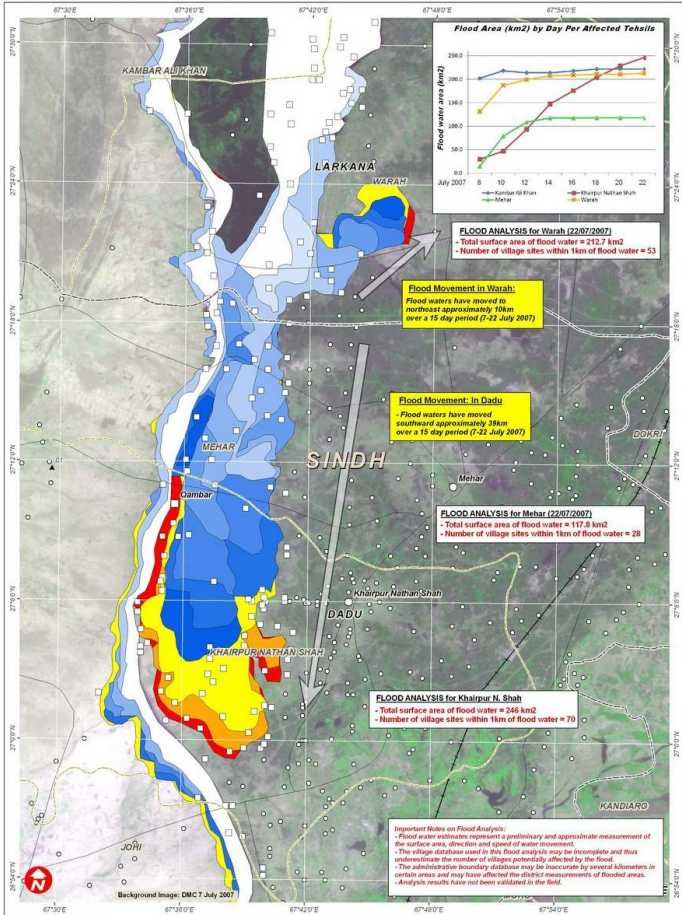
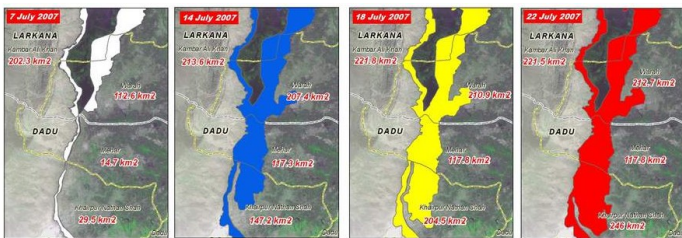
- Arqueología
- Astronomía
- Astrofísica
- Protección civil
- Química
- Ciencias de la Tierra
- Finanzas
- Fusión
- Geofísica
- Física de Altas energías
- Ciencias Naturales
- Multimedia
- Ciencia de materiales ...





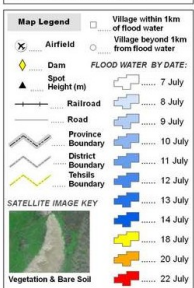
# Aplicaciones prácticas : usando la Grid

Procesado ultra-rápido  
de imágenes satélite en el  
caso de desastres naturales



**Map Information**  
 The main purpose of this satellite detected flood water changes during a 15 day period (7-22 July 2007) over the affected district of Dadu, Sindh Province, Pakistan. Flood estimates represent a preliminary and approximate measurement of the surface area, direction and speed of water movement. This analysis was made using MODIS Aqua 250m imagery validated in the field. Please read the additional analysis notes on the bottom of this map. This map was produced through a joint collaboration among UNOSAT, WFP and IFRC.

The depiction and use of boundaries, geographic names and related data herein are not intended to be endorsed nor do they imply official endorsement or acceptance by the United Nations. This map was produced by the United Nations Institute for Training and Research (UNITAR) Operational Satellite Applications Programme (UNOSAT). UNOSAT provides satellite imagery and related geospatial information to UN Member States and related development agencies and their implementing partners.



**Map Scale for A3: 1:250,000 (Main Frame)**

0 1 2.5 5 7.5 10 Kilometers

Background Image: DMC 7 July 2007

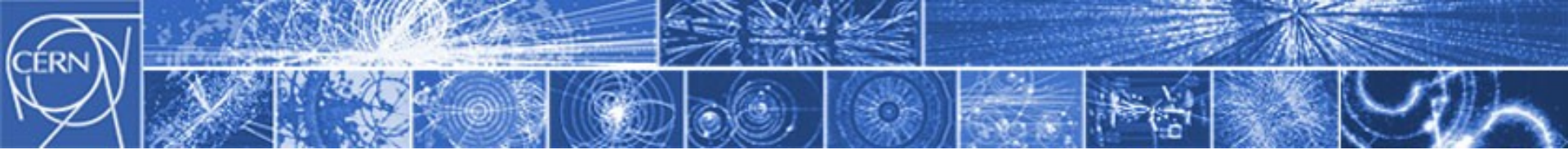
**Map Scale for A3: 1:250,000 (Main Frame)**

0 1 2.5 5 7.5 10 Kilometers

Background Image: DMC 7 July 2007

Copyright © 2007 UNOSAT  
 UNOSAT is a joint venture between WFP and IFRC.  
 UNOSAT is a joint venture between WFP and IFRC.  
 UNOSAT is a joint venture between WFP and IFRC.





# Sin contar... algunos premios Nobel!



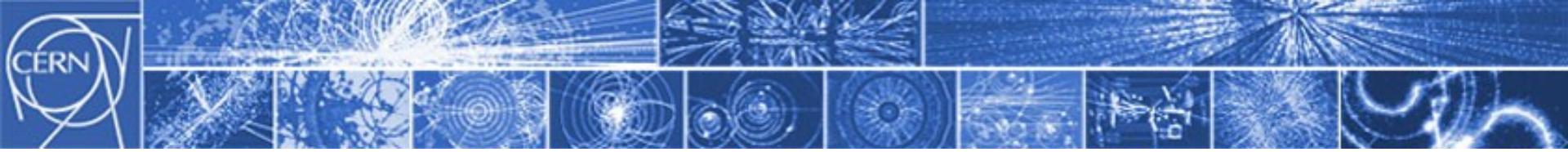
**Carlo Rubbia**  
*(with Simon van der Meer)*

*“for their decisive contributions to the large project, which led to the discovery of the field particles W and Z, communicators of weak interaction”*



**George Charpak**

*“for his invention and development of particle detectors, in particular the multiwire proportional chamber”*



# La Información es Poder...

# Y comprenderla entonces?