

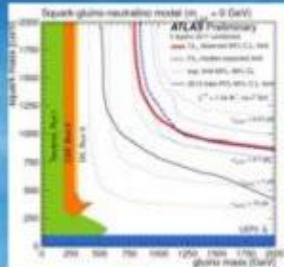
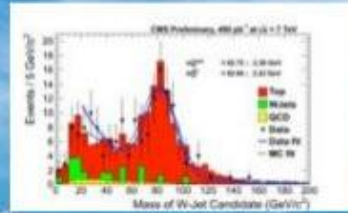
Computación y procesamiento de datos en Física de Partículas

José M. Hernández (CIEMAT, Madrid)

Curso de introducción a la física de
partículas para profesores IES
CIEMAT, Madrid, 9 Marzo 2015

Experimentación en Física de Partículas

Physics
Analysis



Simulation



Reconstruction



R&D

Commissioning



Magnets

Trigger
DAQ

Computing



Installation
Construction



LHC



Esfuerzo global, éxito global

Global Effort → Global Success

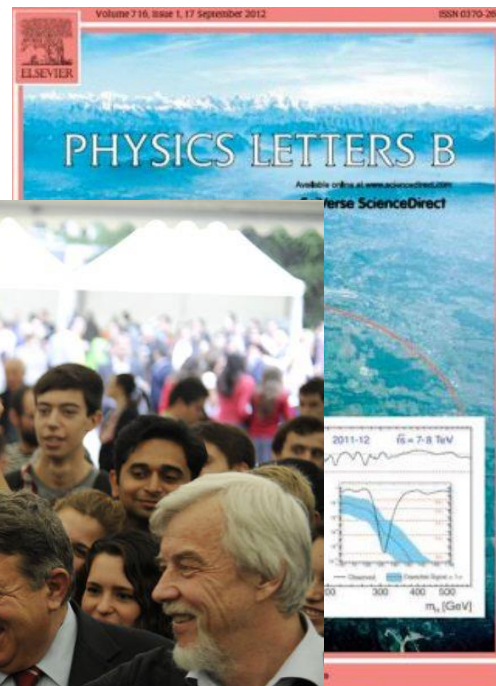
July 4, 2012

Results today only possible due to extraordinary performance of accelerators

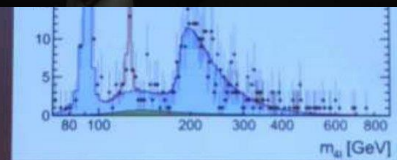
Observation of a Higgs Boson

Historic Milestone

Global Implications

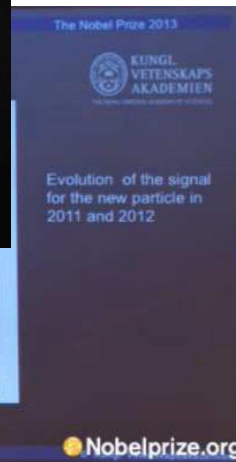


Prince of Asturias Award for Technical & Scientific Research



<https://wiki.cern.ch/wiki/birvine/CMSPublic/Hig13002TWiki>

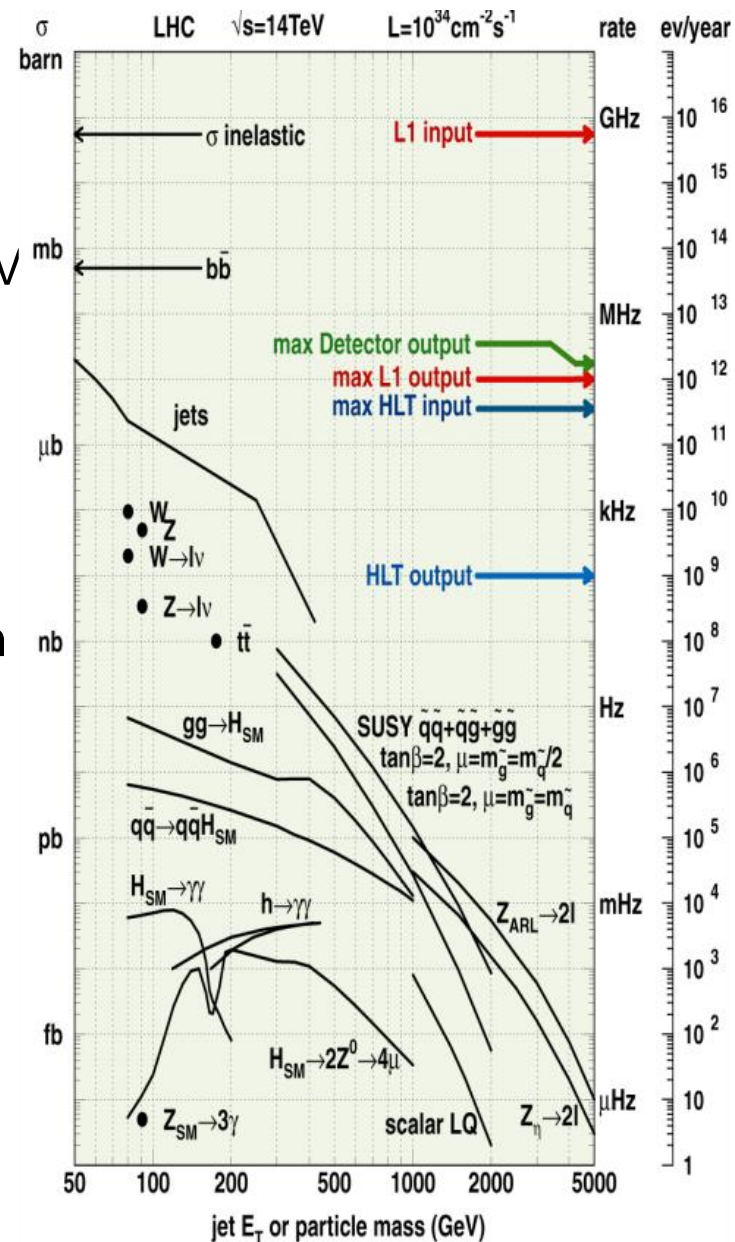
Ogla Bülmel | 8 October 2012



Avalancha de colisiones en LHC

- Baja probabilidad de producción de nueva Física (p.e. bosón de Higgs)
 - A mayor energía de colisión mayor probabilidad de producción: 7,8 \rightarrow 14 TeV
- Tasa de colisión lo más alta posible
 - Flujo de protones (luminosidad instantánea)
 - ~ 3000 paquetes de $\sim 10^{11}$ protones, sección transversal $\sim 10 \mu\text{m}$, que se cruzan cada 25 nanosegundos (40 MHz)
 - ~ 25 colisiones por cruce de haces
- Tasa colisión LHC: ~ 1 GHz
- Tasa producción nueva Física: $\sim \text{mHz}$

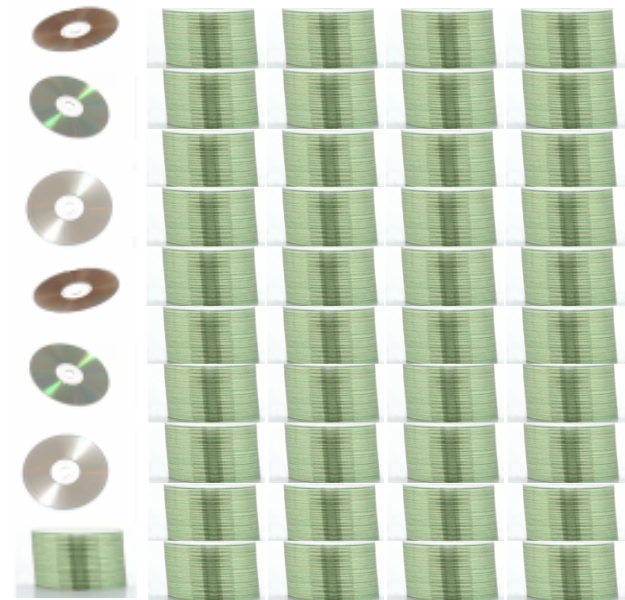
Se necesita filtrar 1 de cada
1.000.000.000.000 colisiones



Volumen de datos

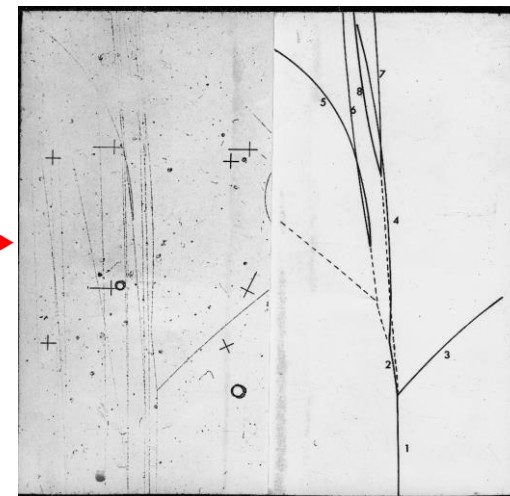
- Si se almacenaran todas las colisiones el volumen de datos sería gigantesco
 - Tamaño registro promedio: ~ 1 Mbyte
 - Número registros/s = 40 MHz
- $\sim 1\text{MB}/\text{registro} \times 40 \times 10^6/\text{s} = 40 \text{ TB/s}$
(0.4 Zettabyte/año)
- Si se guardaran estos datos serían necesarios 60.000 CDs cada segundo

Una pila de 70 metros cada segundo!
- Se necesita filtrar las colisiones interesantes en tiempo real a una tasa manejable
→ “Trigger”

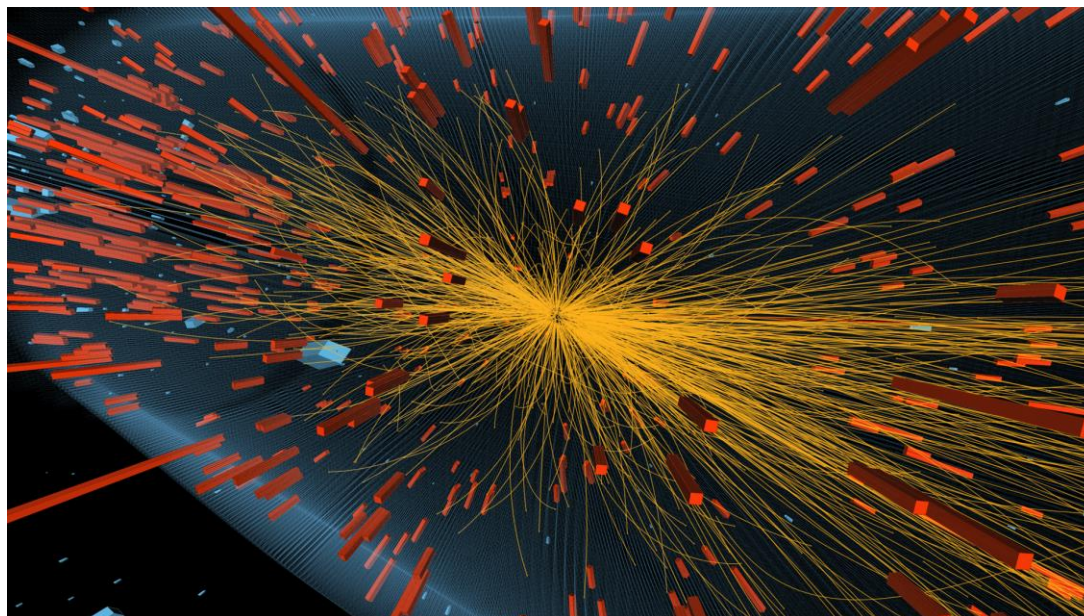


The Trigger

- In the old days the trigger was a manual operation ...

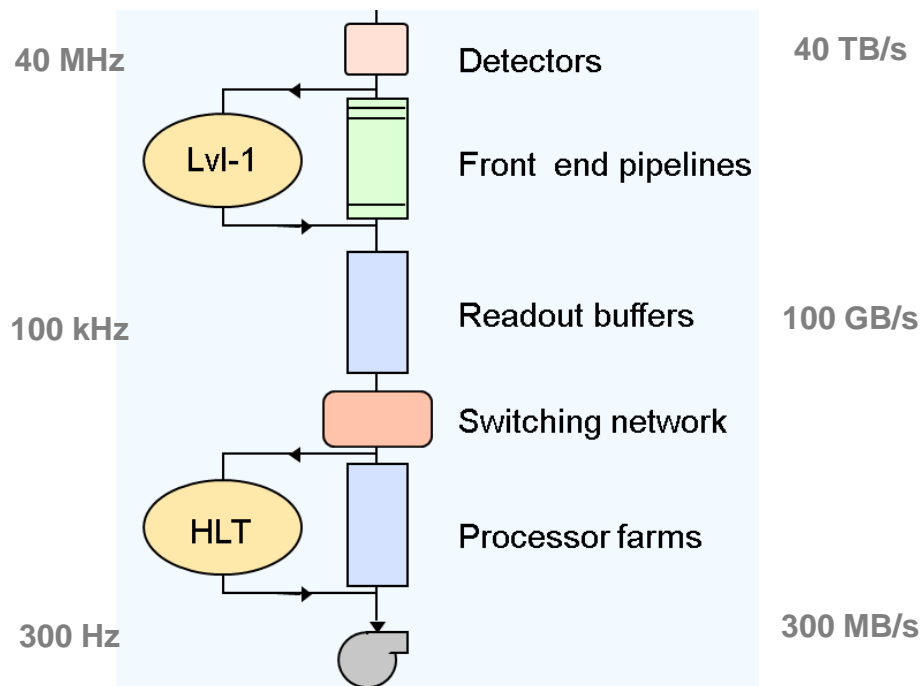


- At LHC one has to filter events like this (1 in 10^{12})

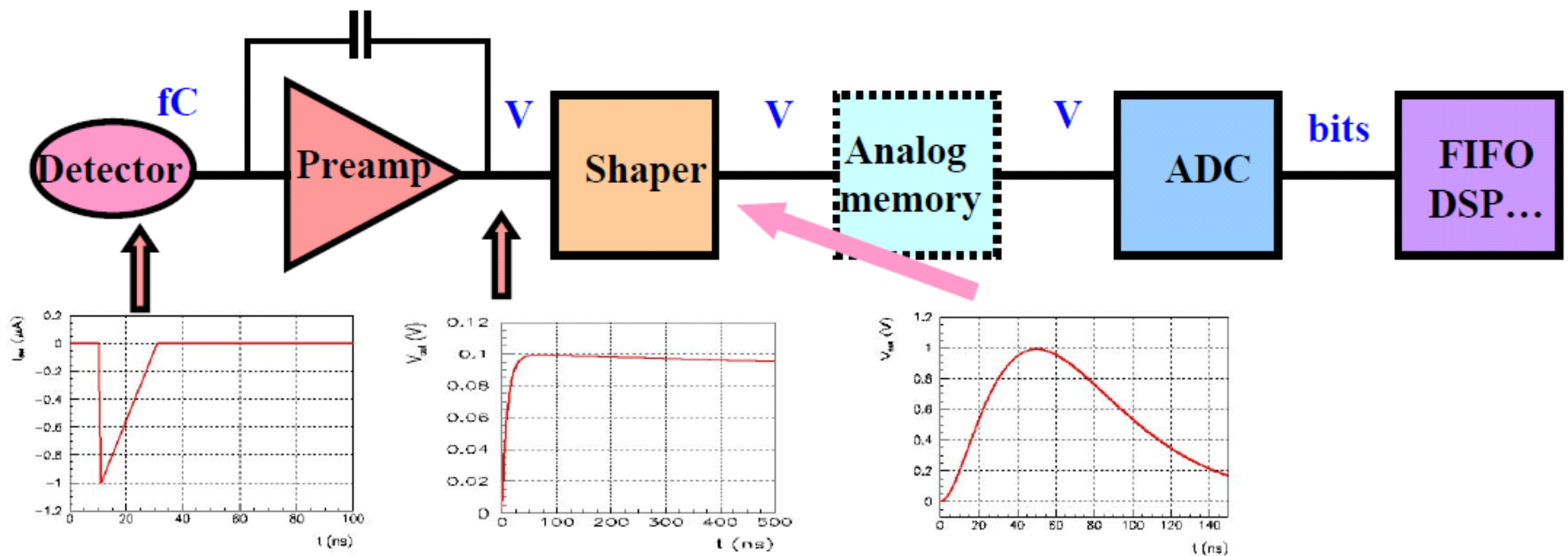


Trigger system

- Nowadays we use a very complex electronic system
- Multi-level, data buffering, parallel processing
- **First Level Trigger**
 - Specialized hardware processors
 - Limited information, simple algorithms
 - 40 MHz \rightarrow 100 kHz
 - 3.2 μ s latency (128 pipeline),
~100 GB/s throughput
- **High Level Trigger** (software)
 - Linux PC farm
 - Flexible software algorithms
 - 100 kHz \rightarrow 300 Hz
 - ~50 ms latency (5000 processors)



Procesado de las señales



- Detección de electrones de ionización/fotoeléctricos: señales muy débiles → **amplificación**
- **Moldeado, Discriminación, Digitalización**
- Medida de la amplitud de la señal (proporcional a la energía de la partícula depositada),
- Media del tiempo transcurrido en detectar la señal (reconstrucción de la posición espacial)

Electrónica del sistema de adquisición de datos



Avalancha de datos

- Tasa aceptación Trigger: ~ 300 Hz
- Volumen de datos anual:
1 MB/registro \times
300 registros/s \times
 10^7 s/año =
3.000.000 Gbytes/año =

3 Pbyte/año



¿Cuánto ocupa un PetaByte?

Ya hemos visto que si lo quisiéramos guardar en CDs, tendríamos una pila de varios km

... evidentemente, tenemos que encontrar otra solución

Una opción que tenemos es almacenarlo en **cinta magnética**

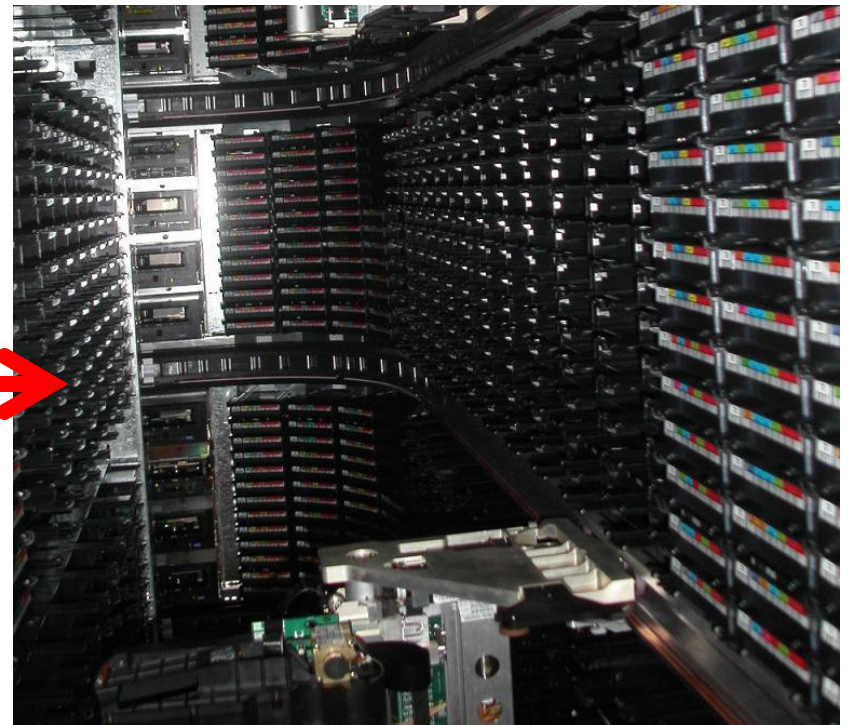


Hoy en día en una cinta de unos 10 x 10 x 2 cm se pueden almacenar hasta 5000 GigaBytes de datos



Almacenamiento en cinta magnética

Para manejar estas cintas de forma eficiente usamos una librería automatizada en la que un **brazo robótico** localiza y accede a los datos guardados



La cinta magnética permite una alta densidad de información ... **pero**, para volver a leer los datos tenemos que **“rebobinar”**

Almacenamiento en discos duros

Otra opción para almacenar los datos es usar discos duros



Un disco duro almacena información en pequeñas áreas que **se magnetizan** positiva o negativamente

Cada una de estas áreas representa un **"bit"** de información: los famosos 0 y 1

La información se lee con un cabezal que escanea la superficie del disco y detecta las variaciones magnéticas

1956: El primer disco duro

La tecnología de los discos duros ha evolucionado de forma impresionante desde que aparecieron los primeros modelos, a mediados de los 50



¡4,4 MegaBytes!
→ 1 archivo mp3...
y de baja calidad!

Año 1956, IBM 305 RAMAC

Física aplicada: Nobel 2007

Fue para los descubridores del fenómeno de la **Magnetoresistencia Gigante**

Que permite utilizar cabezales de discos duros de 50 millonésimas de milímetro

2000 veces más delgado que un pelo humano

El cabezal de un disco duro es extremadamente rápido y preciso

Equivale a un avión supersónico volando a un metro del suelo, y contando las hojas de césped que sobrevuela



¿Cuánto ocupa un PetaByte?

Hoy en día, un disco duro puede tener una capacidad de 4000 GigaBytes = 4 TeraByte

Los discos nos permiten acceder a cualquier parte de los datos en todo momento, sin rebobinar

pero consumen electricidad y **disipan calor**



~200 TB/fileserver
4 TB/disk, 10GE uplinks

¿Cuánto ocupa un PetaByte?

En un armario como éste, llenos de discos, nos puede caber un PetaByte



Así que el espacio físico que ocupa un PetaByte es similar en disco y cinta

Almacenamiento en cinta y disco

Hoy en día, en cualquier centro informático del LHC encontraremos este paisaje:



Datos en bruto, en cinta

Se leen pocas veces, y siempre de forma secuencial



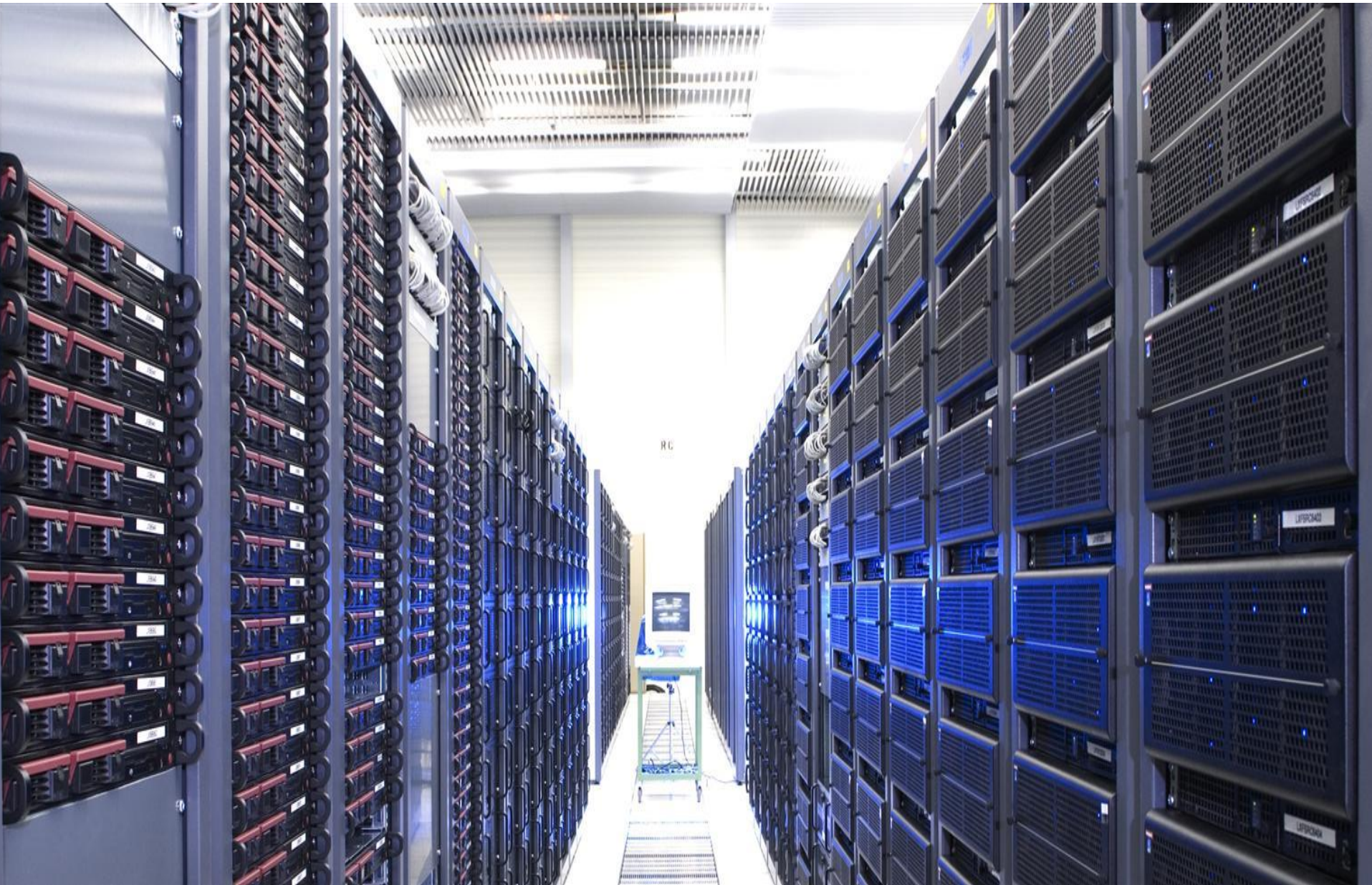
Datos para análisis, en disco

Se leen muchas veces, de forma desordenada

Centro de Computación en el CERN

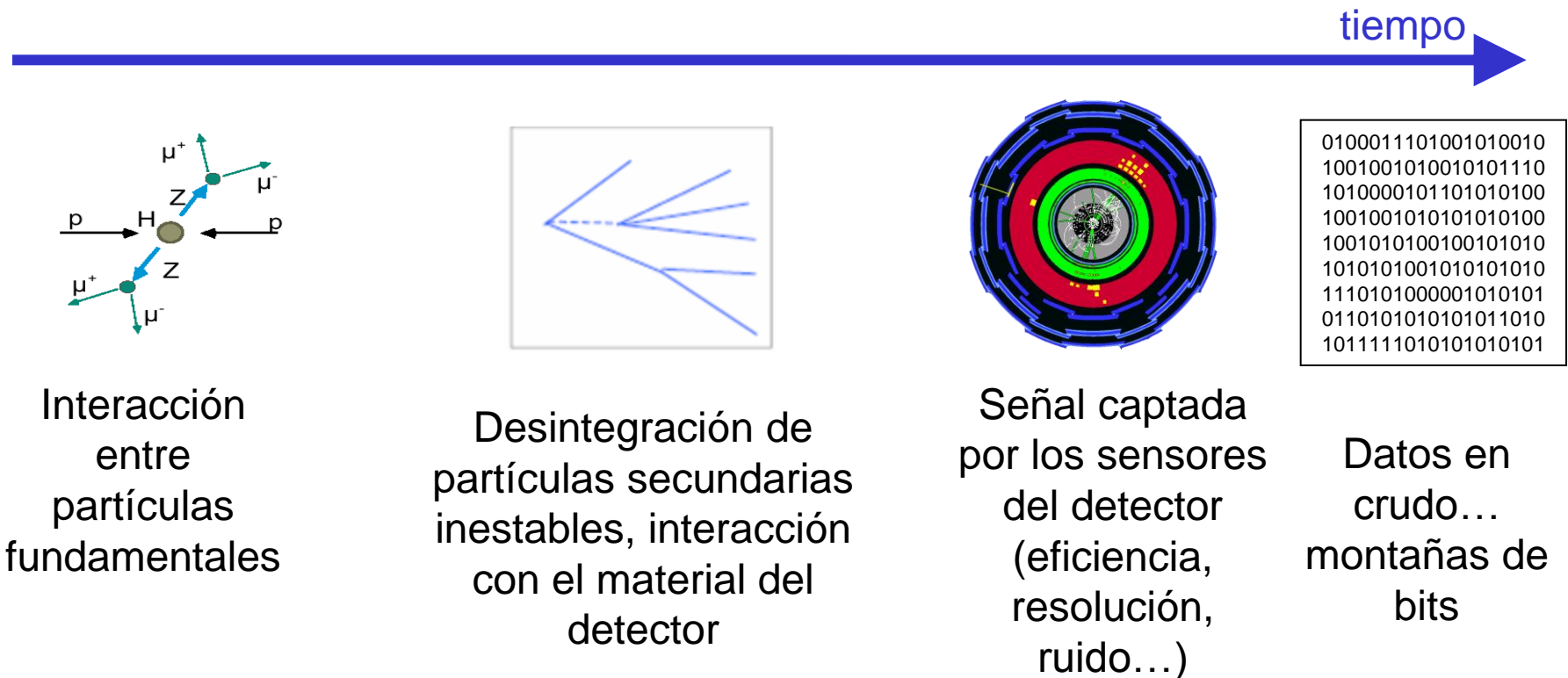


Centro de Computación en el CERN



Los datos

El detector nos lleva de la física hasta los datos

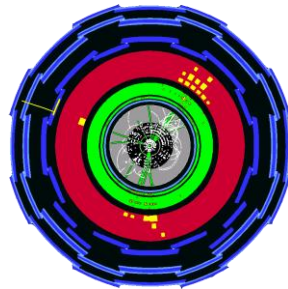


De los datos a la física

Con el **análisis**, tratamos de desandar ese camino

```
0100011101001010010
1001001010010101110
1010000101101010100
1001001010101010100
1001010100100101010
1010101001010101010
1110101000001010101
0110101010101011010
1011111010101010101
```

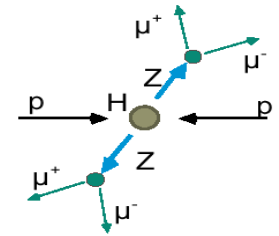
Datos en
bruto
Asociarlos a
cada sensor



Respuesta del
detector
Aplicar
calibración y
alineamiento,



Reconocimiento de
patrones (trazas,
vértices, depósitos
de energía) e
identificación de
partículas



Análisis:
determinar
eficiencia,
resolución,
comparación con
los modelos
teóricos, etc

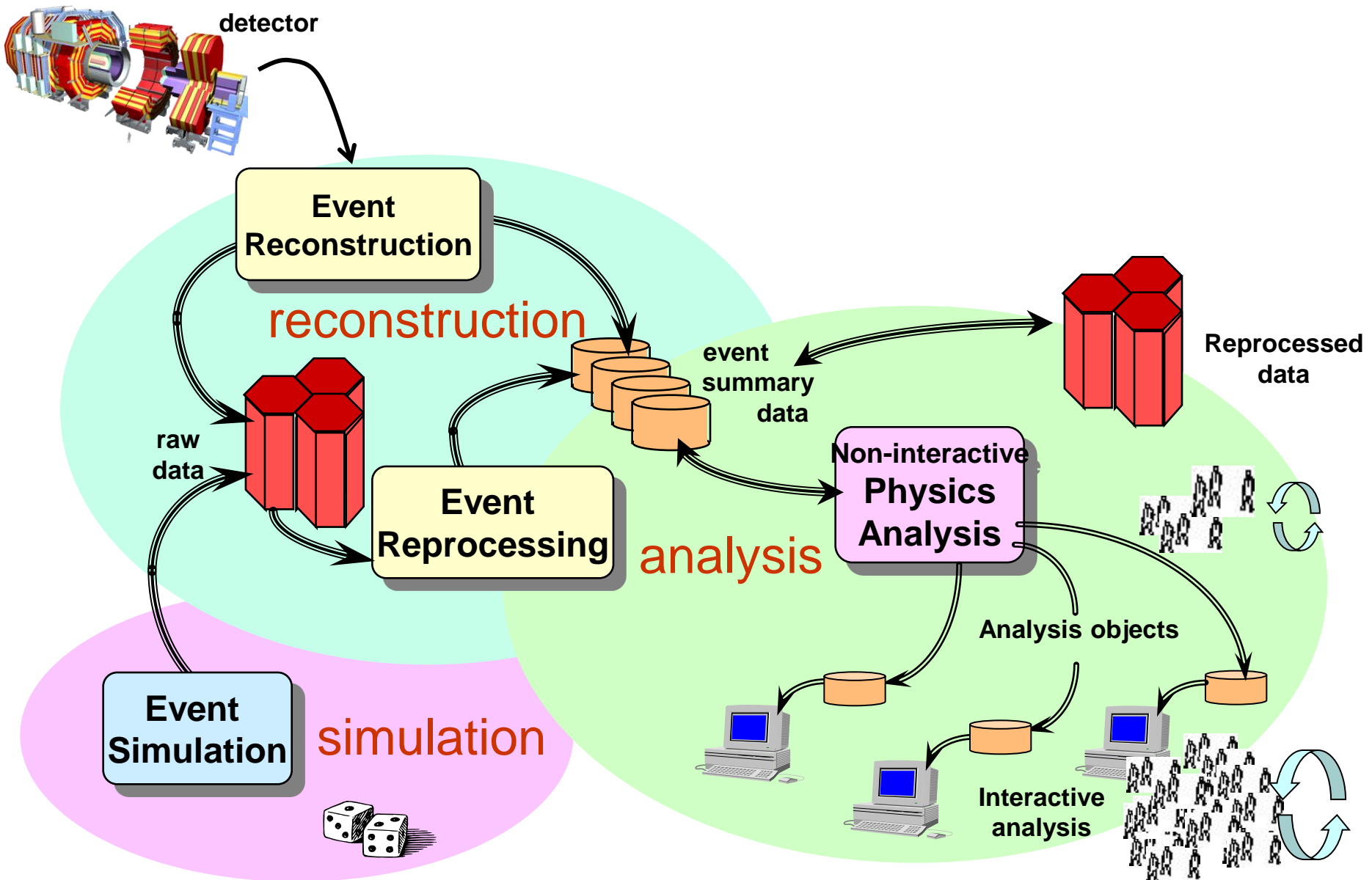
1ª Fase: Reconstrucción

2ª Fase: Análisis

Organización y reducción de datos

- Los datos de las colisiones que pasan el *trigger* se organizan en **datasets**
 - Atendiendo al contenido físico de las colisiones
 - Cada análisis procesa uno o varios datasets determinados
 - Un dataset (~TBs) está compuesto de ficheros de datos (~GBs) que contienen colisiones (~MB) con un determinado formato
- El primer paso del análisis consiste en la **reducción** del número de colisiones o del contenido de las mismas
 - De PBs a TBs
 - Filtrado centralizado, organización en grupos de análisis
- En el siguiente paso cada individuo procesa su selección favorita realizando una selección final
 - De TBs a GBs
- Por último, la selección final se analiza interactivamente con programas especializados de análisis estadístico

Data processing model



El reto de los datos del LHC

- Los datos en bruto deben procesarse (**reconstrucción**) lo cual genera un volumen de datos similar
- Es necesario producir **datos simulados** para entender la respuesta del detector, estudiar predicciones de modelos teóricos, comparar con los datos reales, etc.
- El **análisis** añade un nuevo reto a la gestión de los datos
- En total se producen **~10-20 PB** datos anualmente por experimento

- La reconstrucción, simulación y análisis involucran cálculos muy complejos que requieren de una potencia de cálculo muy grande equivalente a **cientos de miles de PCs**

Servidores de cálculo



quad/six core procs, blades
→ ~500-700 cores/rack



El reto de los datos del LHC

En el LHC, por primera vez en la historia de la física de partículas, ha resultado **imposible** almacenar y analizar los datos en **un único centro** de computación

- LEP, 1989-2000

- Generó en más de una década menos de 1 por mil de los datos de un año de LHC

- Tevatron, 1983-2011

- Generó en toda su vida útil aproximadamente un 25% de los datos de un año de LHC

Divide y vencerás

- El problema del procesamiento y análisis de los datos de LHC se puede dividir y distribuir
 - Cada colisión registrada se puede procesar independientemente
- No es necesario un superordenador
 - Recursos muy caros y de difícil acceso
 - Utilizado para una única aplicación compleja que se puede paralelizar y ejecutar en muchos nodos a la vez y que requiere una comunicación muy rápida entre nodos y un acceso común a la memoria del superordenador → *High Performance Computing*
 - Grandes requisitos de memoria, procesamiento, velocidad
 - Aplicaciones de Meteorología, fusión nuclear, etc

High Throughput Computing

- Requerimientos de potencia de cálculo y memoria para ejecutar una aplicación de procesamiento de datos de LHC son modestos
- Se pueden utilizar ordenadores personales ordinarios
- Lo importante es el resultado agregado de procesar miles de millones de colisiones a través de cientos de miles de trabajos
- High Throughput Computing
- Propuesta para el procesamiento y análisis de datos de LHC
 - Utilicemos los recursos informáticos disponibles en los centros que participan en los experimentos
 - Desarrollemos un sistema para federar esos recursos heterogéneos y dinámicos

Computación Grid

En el LHC, por primera vez en la historia de la física de partículas, se almacenan y se analizan los datos en un entramado de centros de computación interconectados entre sí a través de Internet → **Sistema Computación Grid**

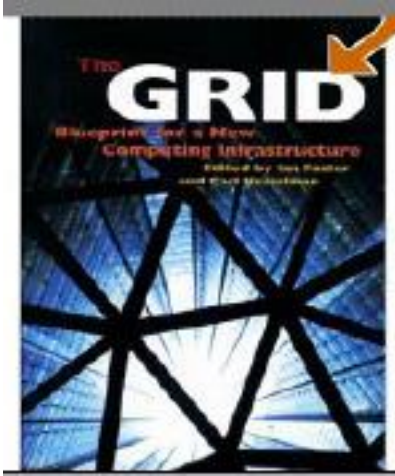
Como un emule (p2p)
gigantesco...



20 años de historia: del WWW al Grid

- 1989:** Tim Berners-Lee propone en el CERN un proyecto de hipertexto global (<http://>)
- 1990:** Windows 3.0 – popularización del PC – millones de personas pueden crear contenido digital
- 1991:** primer sitio web <http://info.cern.ch>
- 1993:** 12 usuarios del 1^{er} navegador, 50 sitios web
- 1995:** Navegador Netscape sale a bolsa – Inicio burbuja tecnológica
- 2001:** explosión de la burbuja .com
 - El mundo queda hiper-conectado: los proyectos de despliegue del Grid a gran escala se hacen factibles
- 2009:** Primeros datos del LHC, analizados por miles de investigadores de todo el mundo a través del **Grid**

“The Grid” (Ian Foster y Carl Kesselman, 1998)



- 1990's boom of accessible computing (PC, better communication networks, Internet, Linux, etc)
- State of computing similar to the development of electricity at the beginning of 1900
- The real revolution of electricity was the possibility to distribute it over a network
- The use of computational services should be as transparent as using a power plug
- Users don't need to know from where the computing power is coming from
- Computing revolution similar to the invention of the Web at CERN



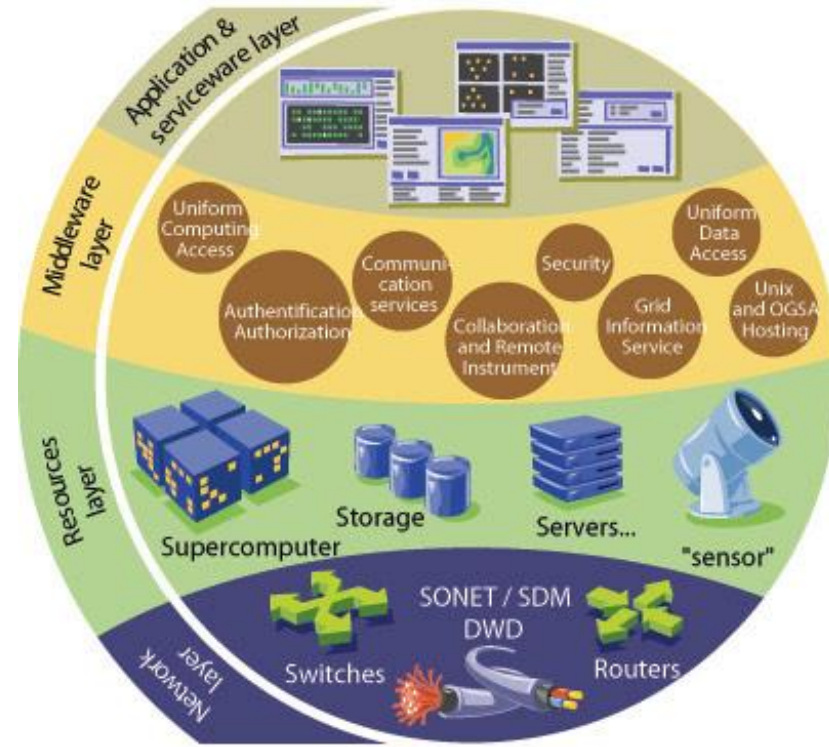
Computación Grid

- El **Grid**, integrando conectividad, computación e información, proporciona una **plataforma virtual para cálculo y gestión de datos**, del mismo modo que la Web integra recursos para formar una plataforma virtual para la información
- El Grid provee la posibilidad de **enlazar dinámicamente recursos heterogéneos** que soportan **ejecuciones a gran escala**, uso intensivo de recursos y **aplicaciones distribuidas**
- El Grid debe proveer **calidades de servicios** no triviales (service level agreements)
- **Complejos sistemas de software y servicios** son necesarios para permitir un **acceso sencillo y seguro** al usuario a través de **protocolos estándar**, para que los recursos sean usados eficientemente, y permitir que las comunidades puedan coordinar sus recursos de forma estable



Arquitectura Grid

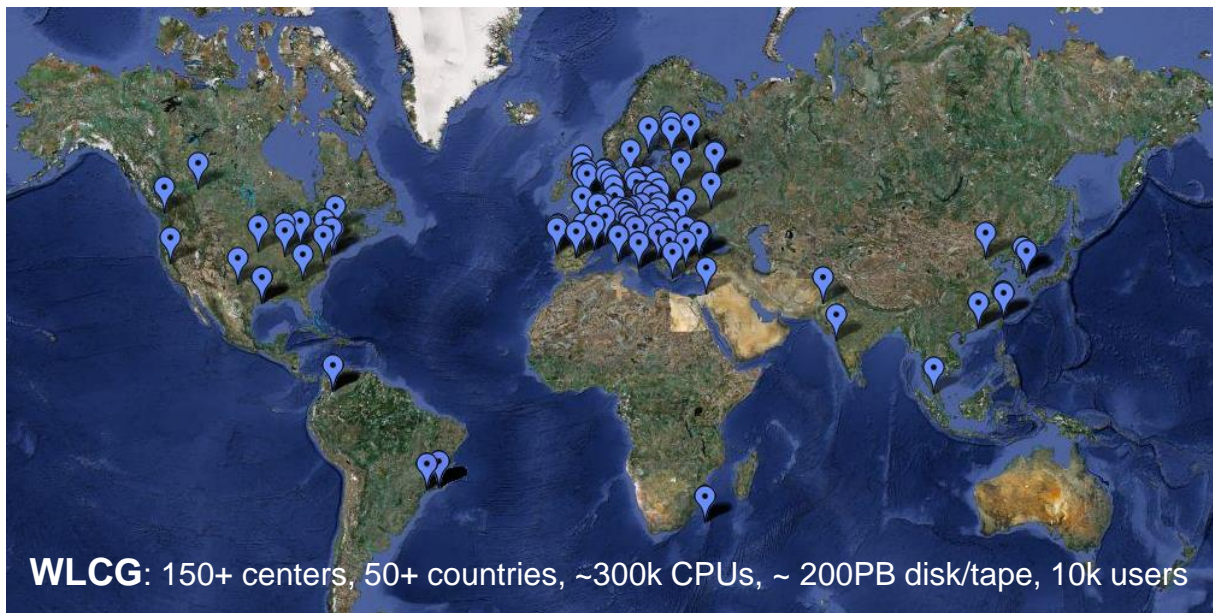
- **Recursos**
 - Computadores, almacenamiento, redes de comunicaciones
 - Heterogéneos, distribuidos geográficamente, dinámicos
- **Middleware**
 - Software que permite conectar y coordinar los recursos
 - Servicios básicos de información, seguridad, gestión de datos y ejecución de tareas de computación, monitorización
- **Aplicaciones**
 - Interacción del usuario con el Grid



El Grid de computación de LHC (WLCG)

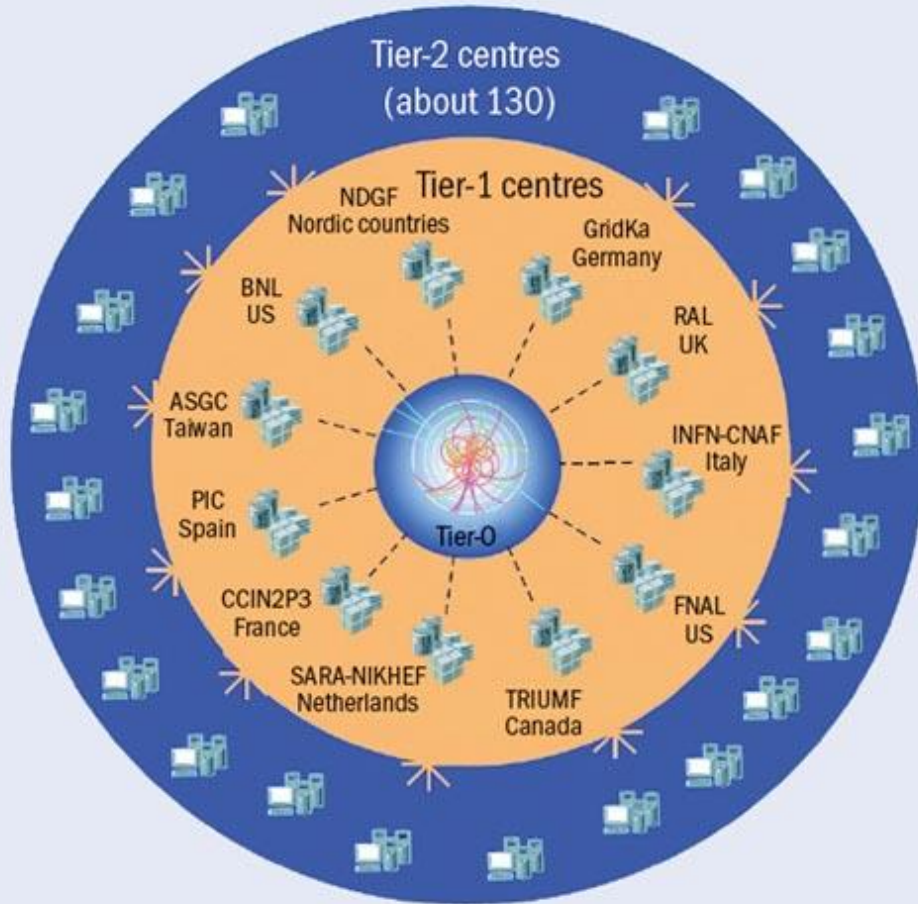
El Grid de LHC es la infraestructura informática que nos permite

- conectar todos los centros de computación de LHC e integrarlos en un único “super-ordenador”
- hacer accesibles los recursos a miles de científicos que a su vez están distribuidos por el mundo

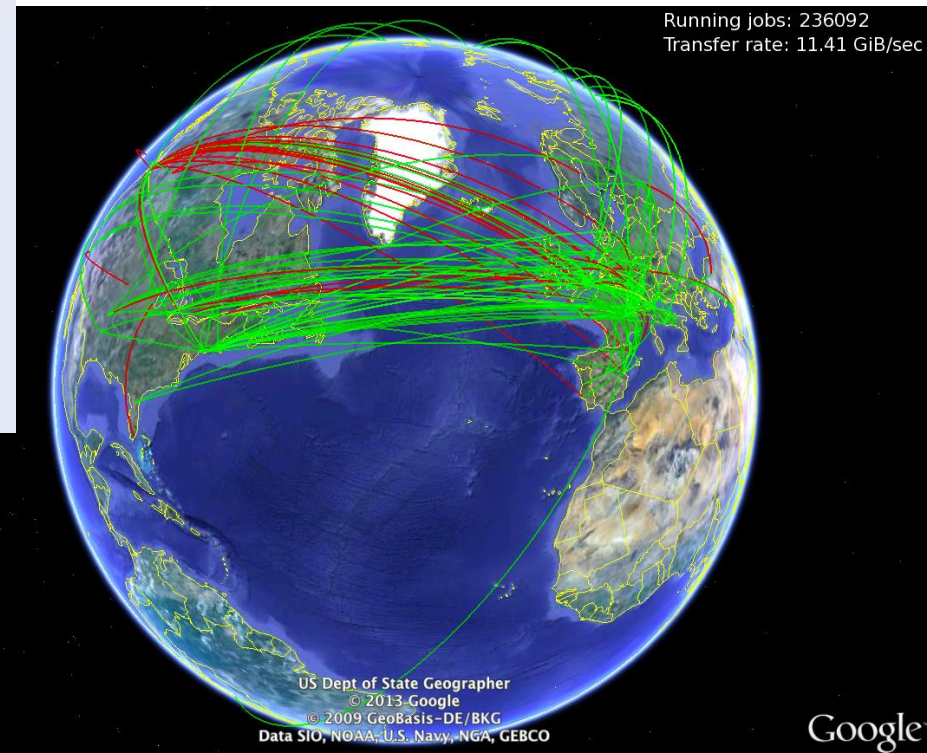


**Estamos ante una auténtica globalización
de los datos del LHC**

Topología de WLCG



WLCG: 150+ centers,
50+ countries, ~300k CPUs,
~ 200PB disk/tape, 10k users



Distribución de datos en tiempo real

Mientras se toman datos en el CERN (**Tier-0**), éstos se envían a 11 centros principales de todo el mundo: llamados **Tier-1**

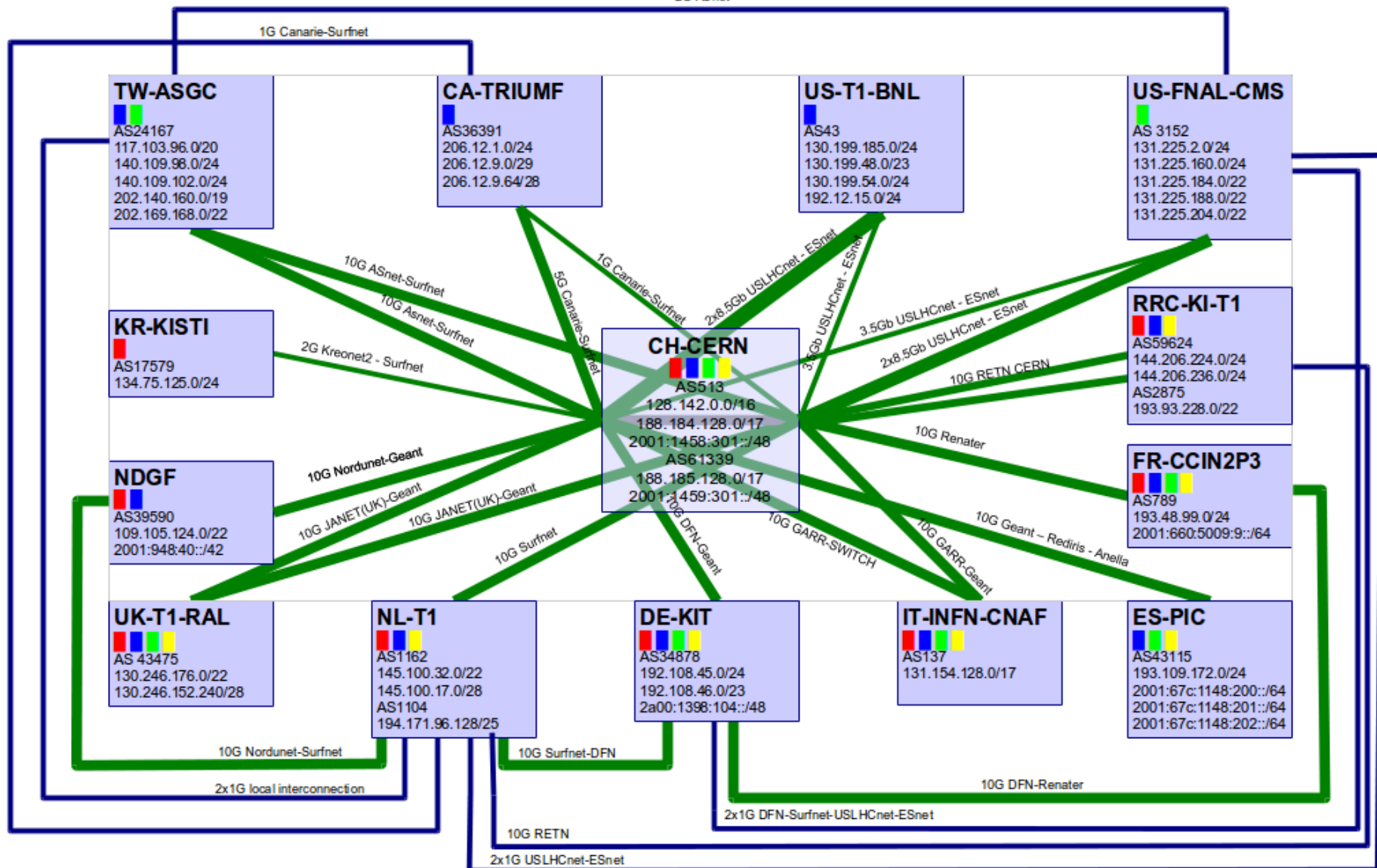


Estos custodiarán la copia de seguridad de los datos en bruto, y se encargarán de la reconstrucción

La conexión entre el CERN y los Tier-1 se realiza a través de enlaces de red privada a una velocidad de **10.000 Megabits por segundo - más de 1000 líneas ADSL**

LHCOPN

2G ASnet



TW-ASGC
 AS24167
 117.103.96.0/20
 140.109.98.0/24
 140.109.102.0/24
 202.140.160.0/19
 202.169.168.0/22

CA-TRIUMF
 AS36391
 206.12.1.0/24
 206.12.9.0/29
 206.12.9.64/28

US-T1-BNL
 AS43
 130.199.185.0/24
 130.199.48.0/23
 130.199.54.0/24
 192.12.15.0/24

US-FNAL-CMS
 AS 3152
 131.225.2.0/24
 131.225.160.0/24
 131.225.184.0/22
 131.225.188.0/22
 131.225.204.0/22

KR-KISTI
 AS17579
 134.75.125.0/24

CH-CERN
 AS513
 128.142.0.0/16
 188.184.128.0/17
 2001:1458:301::/48
 AS61339
 188.185.128.0/17
 2001:1459:301::/48

RRC-KI-T1
 AS59624
 144.206.224.0/24
 144.206.236.0/24
 AS2875
 193.93.228.0/22

NDGF
 AS39590
 109.105.124.0/22
 2001:948:40::/42

FR-CCIN2P3
 AS789
 193.48.99.0/24
 2001:660:5009:9::/64

UK-T1-RAL
 AS 43475
 130.246.176.0/22
 130.246.152.240/28

NL-T1
 AS1162
 145.100.32.0/22
 145.100.17.0/28
 AS1104
 194.171.96.128/25

DE-KIT
 AS34878
 192.108.45.0/24
 192.108.46.0/23
 2a00:1398:104::/48

IT-INFN-CNAF
 AS137
 131.154.128.0/17

ES-PIC
 AS43115
 193.109.172.0/24
 2001:67c:1148:200::/64
 2001:67c:1148:201::/64
 2001:67c:1148:202::/64

— T0-T1 and T1-T1 traffic
— T1-T1 traffic only
 Not deployed yet
 (thick) >=10Gbps
 (thin) <10Gbps
■ = Alice ■ = Atlas
■ = CMS ■ = LHCb
 p2p prefix: 192.16.166.0/24 - 2001:1458:302::/48
 edoardo.martelli@cern.ch 20140620

PIC Tier-1 (Barcelona)

El centro Tier-1 español:
Puerto de Información
Científica (PIC), Barcelona
Gestionado for CIEMAT e IFAE
4k CPUs, 6 PB disco, 8 PB cinta



Centros Tier-2

Hay más de 150 centros secundarios que forman también parte del sistema.



Tier-2s asociados a Tier-1 PIC

The inset map shows a detailed view of Europe with several Tier-2 centers marked by yellow pins. A red arrow points from the main world map to this inset. A red dashed arrow points from the CERN location in the top right to the Barcelona area. The centers are labeled as follows:

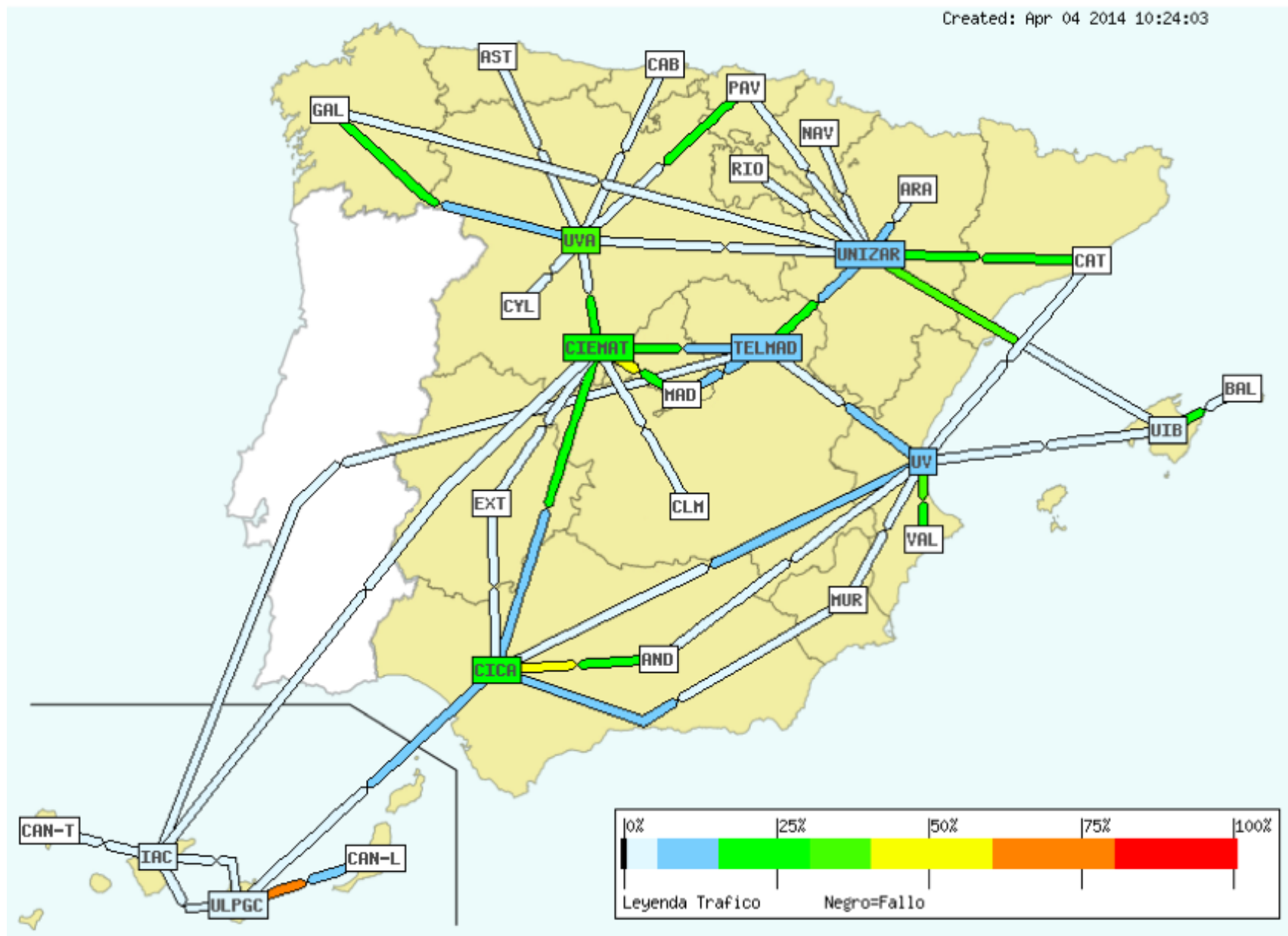
- IFCA [Santander]
- USC [Santiago]
- LIP [Coimbra]
- CIEMAT [Madrid]
- UAM [Madrid]
- IFIC [València]
- LIP [Lisboa]
- IFAE [Barcelona]
- PIC [Barcelona]** (highlighted in green)
- UB [Barcelona]

CERN is indicated by a red dashed arrow in the top right corner.

CMS: CIEMAT&IFCA – LIP_Lisbon&LIP_Coimbra
ATLAS: IFAE&IFIC&UAM – LIP_Lisbon&LIP_Coimbra
LHCb: UB&USC

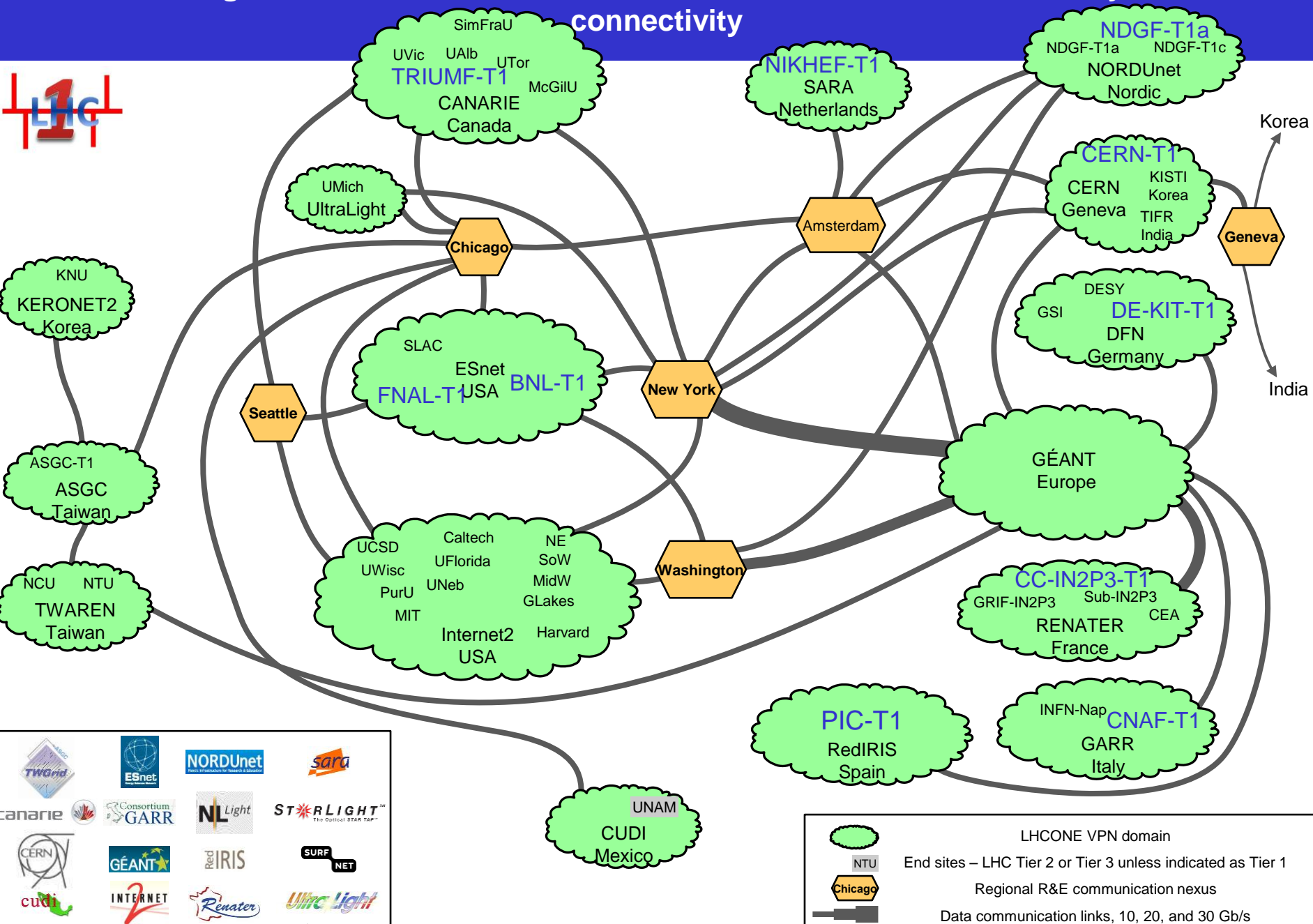
Estos se especializan en la simulación de sucesos y el análisis final de los datos

Rediris: la red académica española



<http://www.rediris.es/conectividad/weathermap/>

LHCONE: A global infrastructure for the LHC Tier1 data center – Tier 2 analysis center connectivity

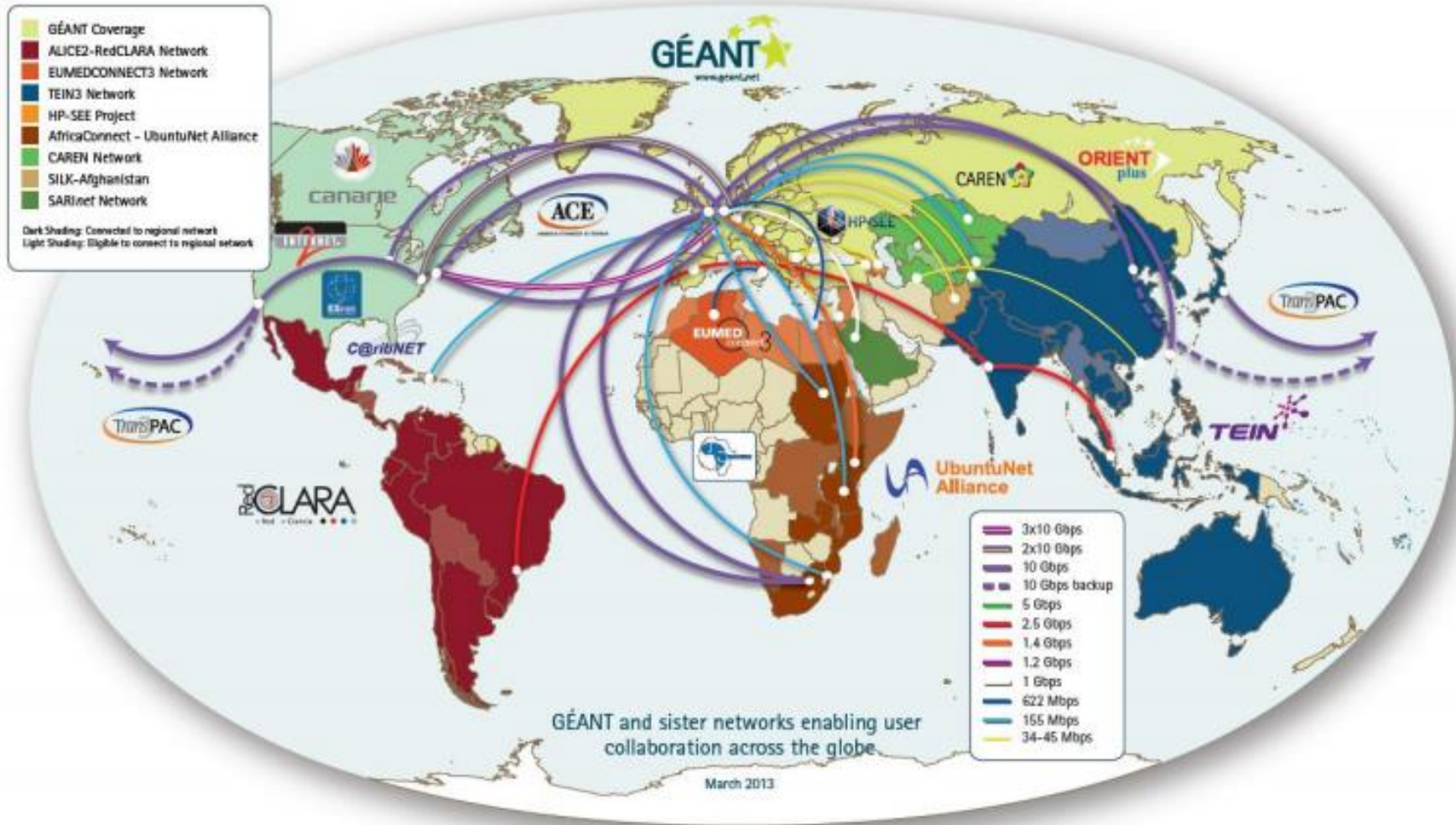


LHCONE VPN domain

- End sites – LHC Tier 2 or Tier 3 unless indicated as Tier 1
- Regional R&E communication nexus
- Data communication links, 10, 20, and 30 Gb/s

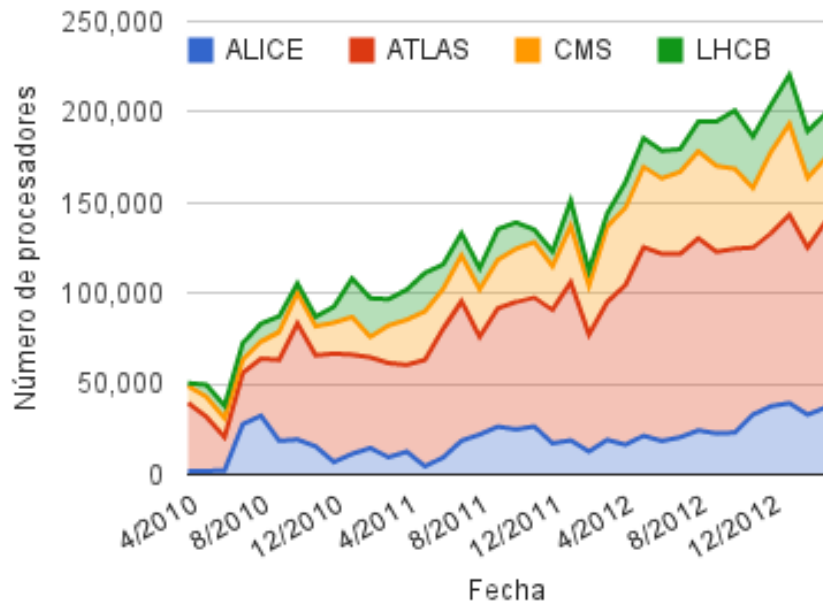
See <http://lhcone.net> for details.

Internet académica a nivel mundial



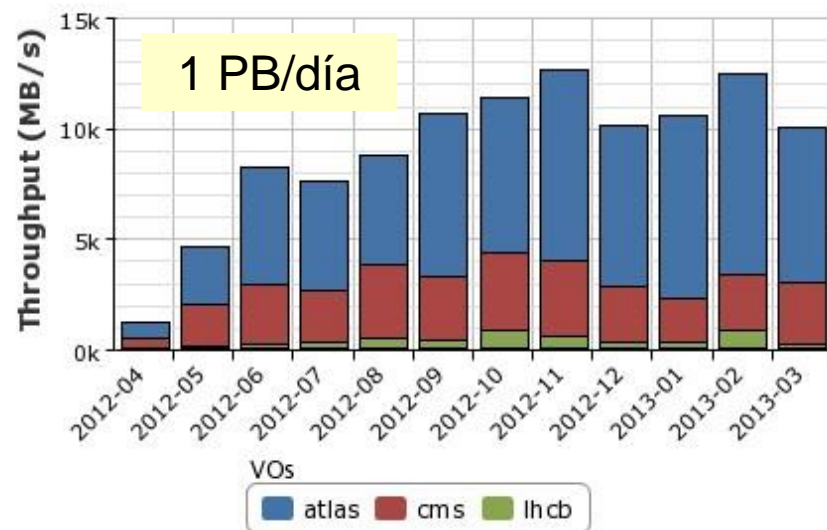
Utilización de WLCG 2010-2012

Utilización de WLCG

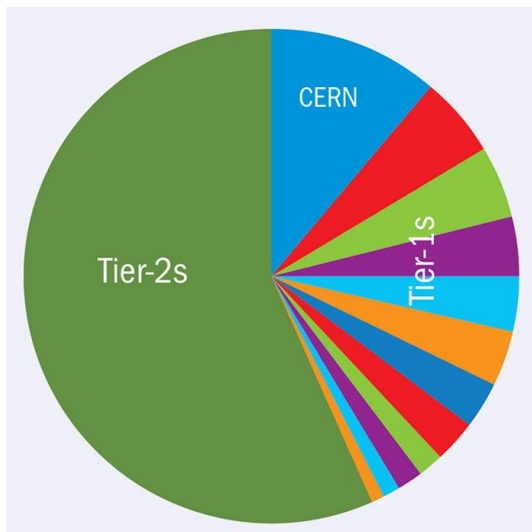
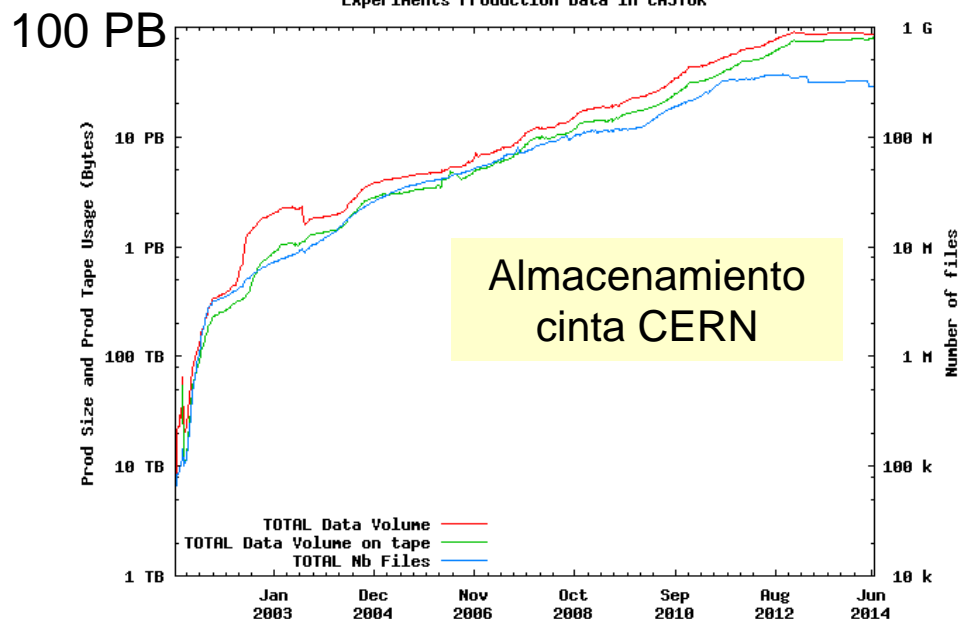


Throughput

2012-04-01 00:00 to 2013-03-31 00:00 UTC



Experiments Production Data in CASTOR



Worldwide LHC Computing Grid

Explosión de datos científicos

La **física de partículas no es la única disciplina** en la que hay una explosión de información

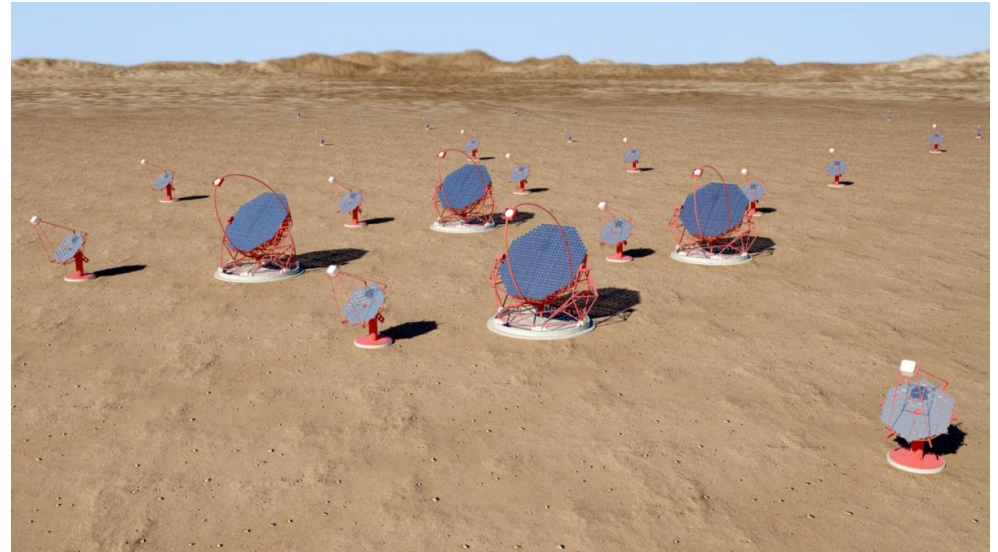
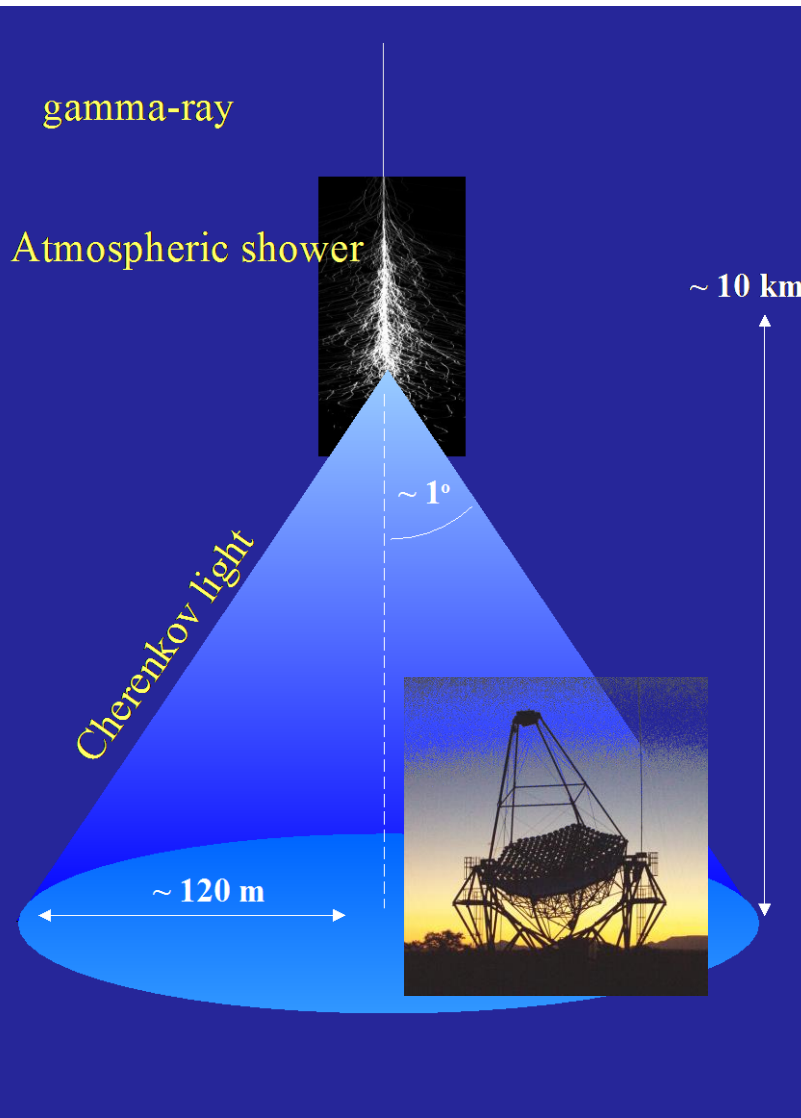
Los instrumentos científicos se digitalizan, y su precisión aumenta

– Crecimiento exponencial de los datos en muchos **otros campos**

En pocos años tendremos telescopios que rastrearán el cielo cada noche, con una precisión sin precedentes

– **10 PetaBytes de imágenes cada año**

The Cerenkov Telescope Array

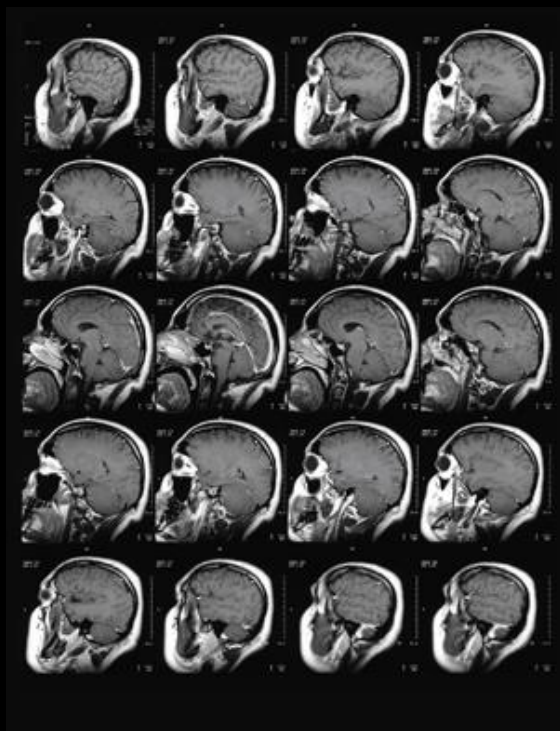


- Grupo de detectores de rayos gamma cósmicos
- 1 sede en hemisferio norte (1 km^2) y 1 sede en hemisferio sure (3 km^2)
- 10-20 PB datos anuales

Explosión de datos científicos

En los hospitales, los **instrumentos de diagnóstico** también se están digitalizando

Se están construyendo bancos de imágenes médicas para investigación que crecerán a razón de **decenas de PetaBytes por año**



¡Muchísimos datos!



Instagram



Google

YouTube

- 300 millones de imágenes subidas a facebook cada día
- 400 millones de tweets enviados al día
- Mil millones de visitas a YouTube / mes
 - Y 100 horas de vídeo subidas por min
- 10.000 millones de móviles conectados para 2020
- Whatsapp procesó 27.000 millones de mensajes en un solo día en 2013



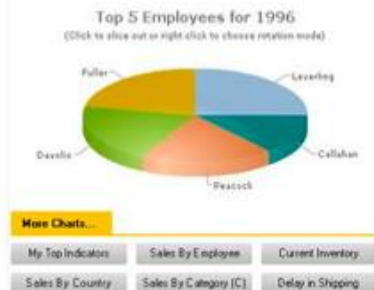
Y desde cualquier parte

- ADSL, wifi, 3G, 4G...
- Móviles, tabletas, gafas (neveras, coches...)
- En 2012, la cantidad global de datos alcanzó los 2,7 ZB



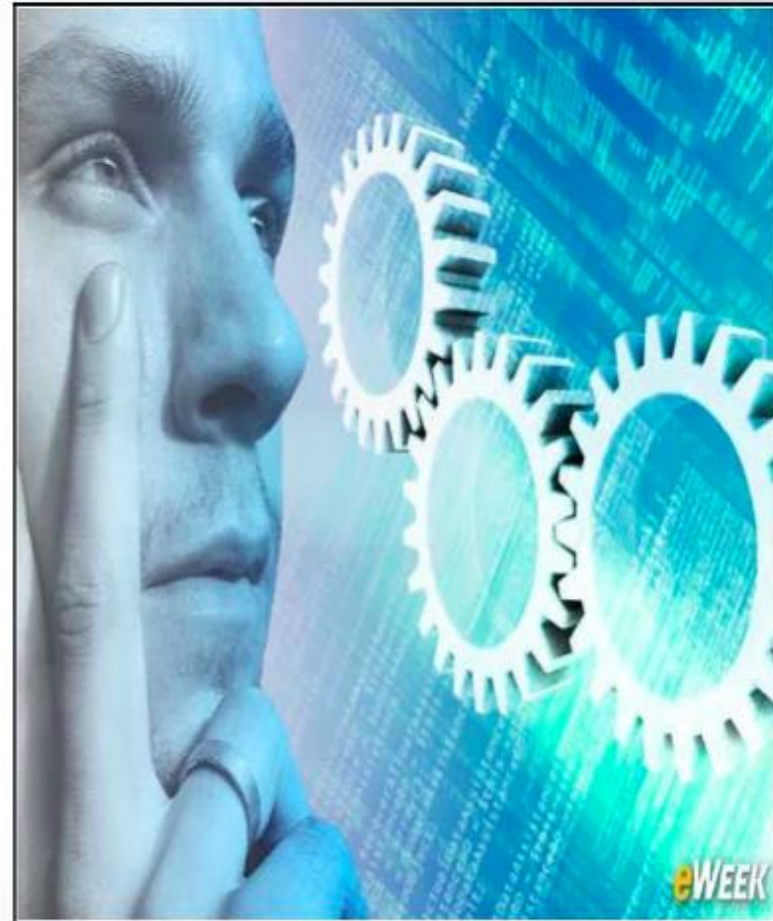
También las empresas

- Datos de clientes y proveedores
- Análisis de mercados, hábitos, tendencias, riesgos
- Simulaciones, diseños
- Quién sabe cuántas cosas más...



¿Qué aportan los datos?

- **Conocimiento**
 - Relaciones y causas
 - Necesidades
 - Oportunidades
- **Transparencia**
 - Mejorar accesibilidad
- **Mejores decisiones**
 - Basadas en datos, no en conjeturas
- **Innovación**
 - Nuevos desarrollos basados en el nuevo conocimiento



Gestión de los datos

- Almacenar
- Acceder
- Mover
- Procesar
- Volver a almacenar
- ...
- *Extraer conocimiento*



Big Data



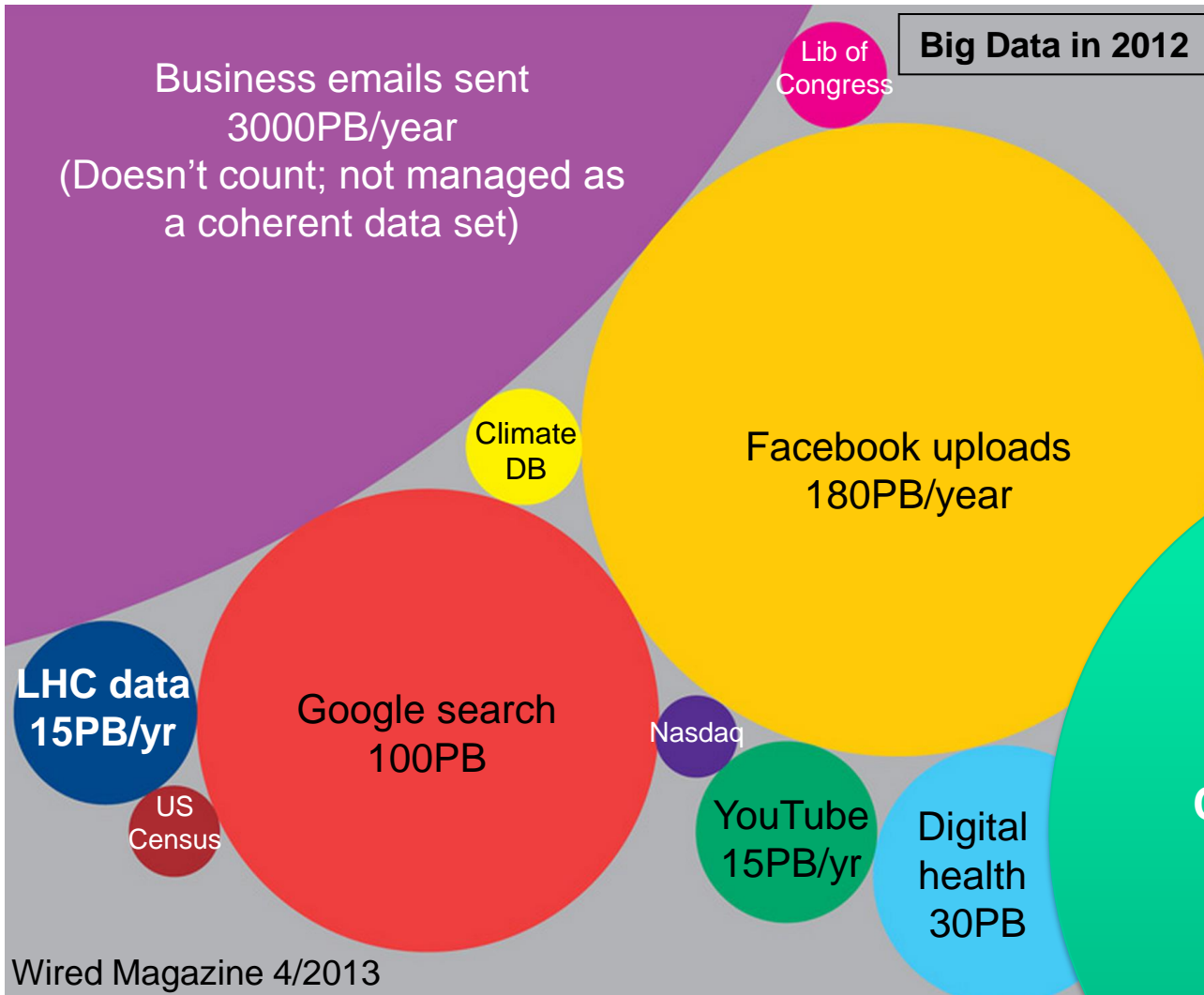
Gestión de los datos

- Where is LHC in Big Data Terms?



Reputed capacity of NSA's new Utah data center: 5000 PB (50-100 MW, \$2 billion)

Big Data in 2012



We are big...

Current LHC data set, all data products: ~300 PB

Wired Magazine 4/2013

¿Computación Grid comercial?

- ¿Es la computación Grid la solución para empresas y usuarios?
- Recursos de computación operados por terceros permiten reducir costes, aumentar fiabilidad y flexibilidad
- Sin embargo el Grid no ha cumplido con el **sueño de transparencia, flexibilidad y facilidad de uso**
 - Operación costosa y complejidad de uso
 - Las aplicaciones requieren generalmente de una configuración específica de los nodos de computación y de servicios específicos
- La arquitectura Grid no es adecuada para el modelo de negocio empresarial (pago por uso)
 - El modelo Grid está basado en proyectos que tienen asignados el uso de una fracción de los recursos
- El Grid no es adecuado para aplicaciones interactivas
 - Latencia significativa en la asignación de recursos

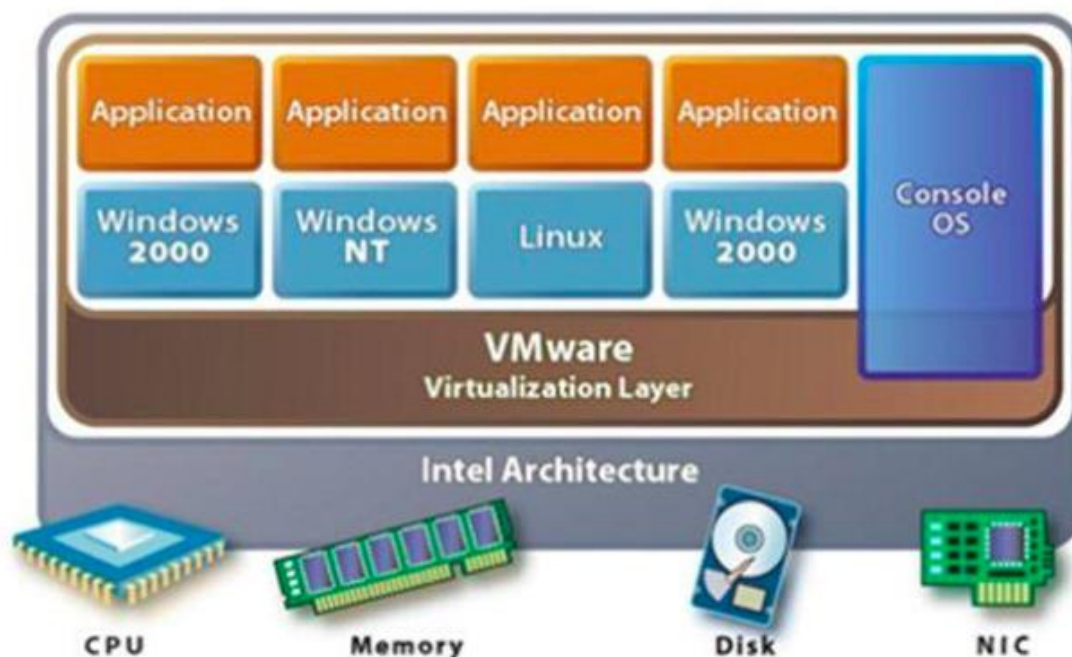
Computación en la nube

- Infraestructura de computación totalmente transparente para el usuario



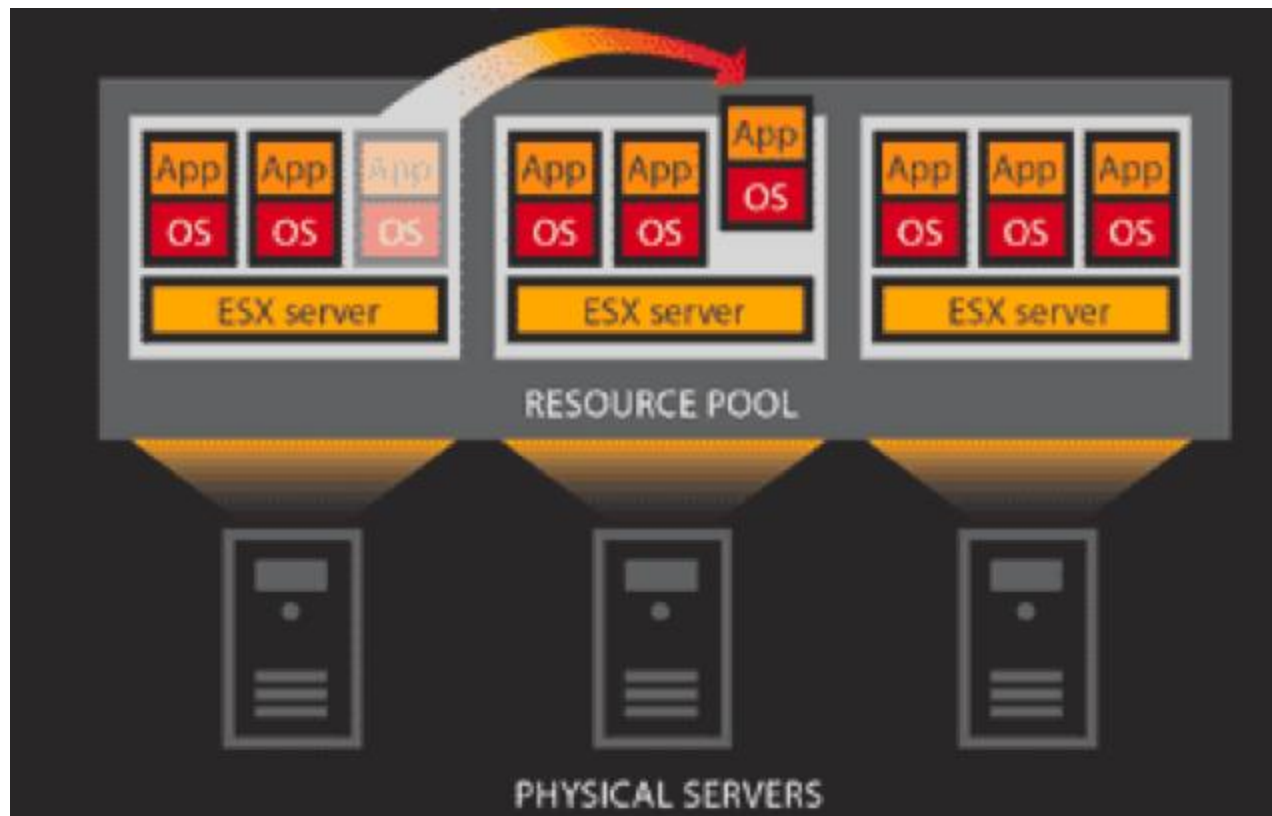
Virtualización

- La tecnología que permite la computación en la nube
- Un ordenador virtual dentro del ordenador físico. O varios!
- La aplicación del usuario es ejecutada dentro de un proceso que suministra un sistema operativo virtual



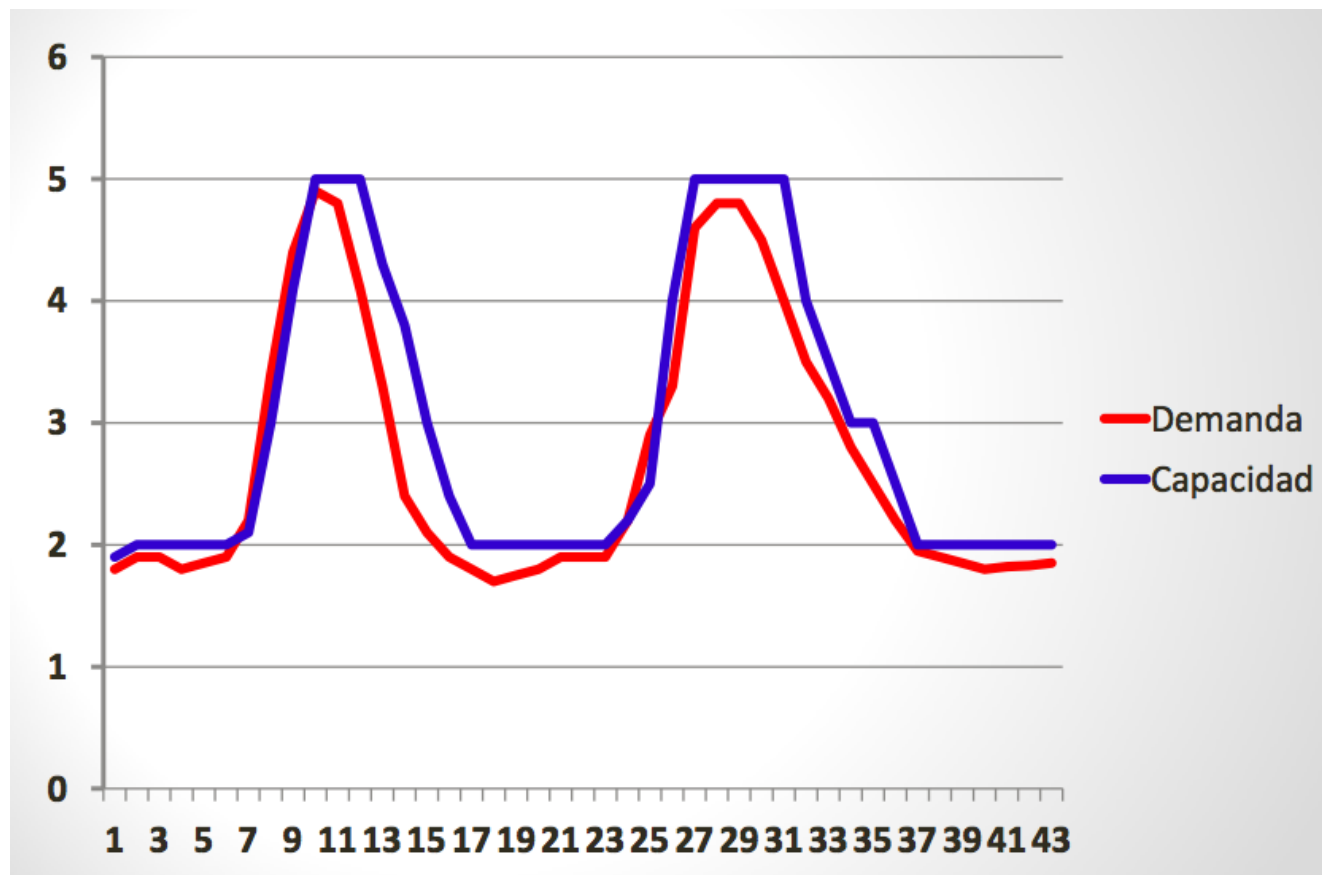
Virtualización y flexibilidad

- La virtualización permite arrancar y mover aplicaciones entre servidores según las necesidades



Virtualización y capacidad bajo demanda

- La capacidad se puede adaptar a la demanda sin más que arrancar o apagar máquinas virtuales



Todo en la nube



Dropbox

Almacenamiento

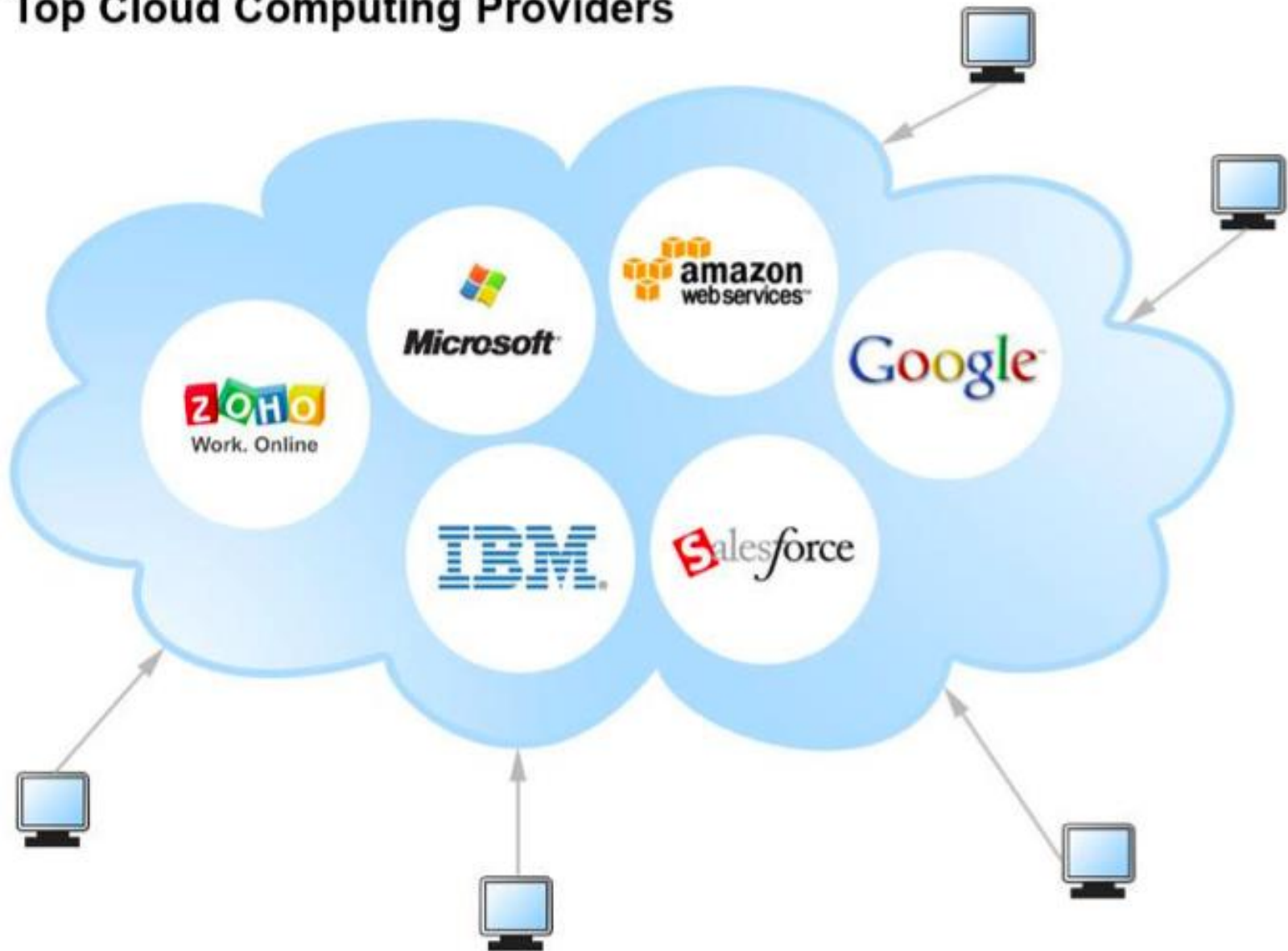


Aplicaciones



Servicios para empresas/ciencia

Top Cloud Computing Providers



El corazón de la nube

