



Ciclo de conferencias sobre física de partículas de la Fundación BBVA

El LHC es la máquina científica que más datos genera, y puede ‘aprender’ a seleccionar cuáles son los importantes, explican Maite Barroso y Pippa Wells

- Maite Barroso, subdirectora del departamento de Tecnologías de la Información del CERN, afirma que en la actual etapa de operaciones del acelerador LHC se espera obtener unos 50 millones de gigabytes por año, casi el doble de datos que en la primera fase. Barroso ha tenido un papel clave en el desarrollo de la red de computación *grid*.
- Pippa Wells, del Departamento de Física del CERN, asegura que “nuestra creciente capacidad de acumular y procesar grandes cantidades de datos afectará a la humanidad de muchas maneras”.
- Esta edición del ciclo ofrece seis sesiones con expertos del CERN e instituciones colaboradoras. Su objetivo es explicar los desafíos presentes y futuros en física de partículas, dar a conocer las tecnologías de los grandes aceleradores y poner de manifiesto los beneficios que aporta la ciencia a la sociedad

Madrid, 13 de mayo de 2016.- “A día de hoy, y en el mundo científico, el acelerador LHC es la máquina que más datos genera”, dice Maite Barroso, subdirectora del departamento de Tecnologías de la Información del CERN. Los números no son pequeños: el Gran Colisionador de Hadrones, LHC por sus siglas en inglés, recaba cada segundo el equivalente a cuarenta millones de imágenes en alta definición, y debe decidir al instante cuáles almacenar para su posterior análisis. Hacer frente a este reto ha exigido desarrollar algoritmos matemáticos específicos, y la mayor red de recursos de computación del planeta. De esta nueva tecnología hablaron ayer en la Fundación BBVA Barroso y su colega Pippa Wells, del Departamento de Física del CERN.

Ambas intervinieron en el ciclo *El CERN reanuda el funcionamiento del LHC y prepara su futuro*, con la conferencia *La computación del CERN al servicio de la ciencia y su impacto en la sociedad*.

En el LHC, partículas subatómicas –en concreto protones y antiprotones- son aceleradas hasta casi la velocidad de la luz dentro de un anillo subterráneo de 27 kilómetros de largo. Las partículas chocan entre sí, y entonces la gran energía que llevan se convierte en nuevas partículas con más masa. Estudiar estas partículas es el objetivo de los físicos, que para ello deben recoger, almacenar y procesar la enorme cantidad de datos generados en las colisiones. Durante la primera etapa de operaciones del LHC “el volumen de datos al año fue de 30 millones de gigabytes, que equivale aproximadamente a 9 millones de películas en alta definición”, señala Barroso.

Durante la etapa actual, en el llamado *Run 2*, el acelerador funciona a energía e intensidad más elevadas, lo que significa más datos: “Esperamos casi el doble de datos que los producidos en el *Run 1*, alrededor de 50 millones de gigabytes por año”, añade Barroso.

Ya desde los inicios del proyecto del LHC quedó claro que tratar con un volumen tan grande de datos exigiría nuevas técnicas y capacidades de computación. El CERN lideró entonces la creación de una nueva red internacional de computadoras que pudieran procesar los datos de forma coordinada, compartiendo recursos de computación. En la actualidad la Worldwide LHC Computing Grid –llamada simplemente ‘grid’-, “es una colaboración global, coordinada por el CERN, que incluye a mas de 170 centros de datos en 42 países, y que hace uso de infraestructuras de grid nacionales e internacionales”, explica Barroso.

El centro de computación del CERN aporta el 30 por ciento de la CPU necesaria y se ocupa del “archivado, reprocesamiento y la distribución rápida y eficaz de los datos a todo el mundo”, prosigue Barroso. “El resto procede de la red de computación articulada por la Worldwide LHC Computing Grid”.

Cada día se ejecutan en la grid una media de 2 millones de tareas, y la cifra sigue creciendo: “En el *Run 2* nos enfrentamos a una mayor capacidad de almacenamiento y de computación para el análisis de los datos”, señala Barroso. Por ello la capacidad de computación del CERN se ha incrementado con un nuevo centro de datos en Budapest y con la construcción de una “nube” privada de computación basada en el proyecto de código abierto Openstack.

Uno de los principales retos es saber qué datos es necesario conservar y procesar y cuáles constituyen en cambio la paja que oculta lo importante. Para ello el primer paso es la reconstrucción, por parte de las computadoras, de las colisiones entre partículas casi en el instante en que se producen, usando complejos algoritmos de sincronización y filtrado. Esta etapa permite a los físicos identificar las partículas creadas y las colisiones más interesantes, que se guardarán para un análisis más detallado.

Un 1% de información relevante

“El filtrado desecha el 99% de los datos”, explica Barroso. “Los algoritmos que se emplean para ello, y su implementación con técnicas basadas en *machine learning* se mejoran continuamente, y aseguran que el cerca del 1% que se guarda es información relevante”.

En este terreno es experta Pippa Wells, que añade: "Usamos sofisticados algoritmos para seleccionar los eventos interesantes. (...) En mi opinión el mayor reto está en asegurarnos de que tenemos los eventos correctos. Una vez los datos se han grabado entonces podemos mejorar nuestro conocimiento sobre la respuesta del detector, y revisar los eventos más de una vez para asegurarnos de que no perdemos nada".

Wells ha comentado: "Las muchas maneras en que nuestra creciente capacidad de acumular y procesar grandes cantidades de datos afectarán a la humanidad. Las técnicas desarrolladas en el CERN nos hacen más capaces de gestionar grandes cantidades de datos y hallar en ellos determinados patrones. Es más difícil predecir qué impacto tendrá la capacidad de recopilar datos de diferentes aspectos de la vida y emplearlos para inferir información sobre las personas y su comportamiento."

Barroso ha subrayado también otro aspecto esencial en la computación de CERN, la seguridad: "Podemos decir que el centro de datos del CERN está continuamente en el punto de mira de ataques informáticos, y esto se extiende al resto de centros de datos de la grid del WLCG. La seguridad es una prioridad, y el objetivo es encontrar el equilibrio entre la libertad de un centro académico y un nivel aceptable de seguridad informática".

Datos biográficos

Maite Barroso estudió ingeniería técnica de telecomunicaciones en la Universidad Politécnica de Madrid porque le gustaban las matemáticas y la física, "y buscaba una carrera que tuviera las dos y con una aplicación directa visible en el día a día, una vertiente práctica", afirma. Se incorporó al CERN en 2001, después de una experiencia en la empresa privada en el campo de la telefonía: "Fue una oportunidad; yo ya estaba trabajando en Ginebra cuando empezaron los proyectos para desarrollar el grid, y me postulé como candidata". Empezaba entonces el primer proyecto de investigación para desarrollar el grid, coordinado por el CERN.

Actualmente es subdirectora del Departamento de Tecnologías de la Información del CERN, departamento que gestiona el procesamiento y almacenamiento de datos y redes de comunicaciones y soporte para todo el programa experimental. Su puesto también incluye I+D en futuras tecnologías en colaboración con la empresa privada y otros centros de investigación a escala mundial.

Desde sus inicios en el CERN ha trabajado en el área de la computación grid como investigadora y coordinadora de los proyectos de I+D financiados por la Unión Europea y liderados por aquella organización: DataGrid, EGEE y el proyecto Worldwide LHC Computing Grid (LCG).

Pippa Wells, física británica, es violinista desde niña, y afirma que su pasión es tocar en orquestas –en los últimos años lo ha hecho en la Orchestre Symphonique Genevois-. También es física experimental de partículas en el CERN, y se ha ocupado de las mejoras del gran detector ATLAS del acelerador de partículas LHC, que jugó un papel fundamental en el descubrimiento del bosón de Higgs en 2012.

ATLAS –como el resto del LHC- deberá ser sometido a constantes actualizaciones en el marco del proyecto que garantiza que el LHC seguirá siendo el mayor acelerador de partículas en el mundo en las próximas décadas. Este proyecto, el HL-LHC (siglas en inglés de Gran Colisionador de Hadrones de Alta Luminosidad) es el más ambicioso en el ámbito de la física de alta energía para la próxima década.

Wells se doctoró en la Universidad de Cambridge en 1990 y fue contratada por el CERN para trabajar en el acelerador precursor del LHC. Uno de sus trabajos consistió en investigar los efectos de las mareas terrestres en el túnel del acelerador. Ya en el LHC, en el equipo de ATLAS, se ocupó de un aspecto clave: determinar de forma precisa la trayectoria de las partículas de materia que se generan en el acelerador, un detalle esencial para identificarlas.

El CERN y la Fundación BBVA

El CERN y la Fundación BBVA iniciaron su colaboración hace dos años, cuando el organismo científico supranacional decidió celebrar en España el 60 aniversario de su creación en colaboración con la Fundación BBVA. El resultado fue el ciclo de conferencias *Los secretos de las partículas. La física fundamental en la vida cotidiana*, celebrado a lo largo de 2014 y clausurado por el todavía director general del CERN, Rolf Heuer. En el segundo ciclo intervino su sucesora, Fabiola Gianotti. Todas las conferencias están disponibles íntegramente en fbbva.es.

La tercera edición que ahora comienza ofrece seis sesiones, en las que participan quince expertos del CERN e instituciones colaboradoras. Su objetivo es promover los desafíos presentes y futuros en el campo de la física de partículas, además de dar a conocer las tecnologías que emplean las grandes infraestructuras científicas y poner de manifiesto los beneficios que aporta la ciencia a la sociedad. El formato del ciclo refleja la estrecha colaboración entre el CERN y las universidades y centros de investigación españoles.

Fundación BBVA

Para más información, póngase en contacto con el Departamento de Comunicación y Relaciones Institucionales de la Fundación BBVA (91 374 52 10; 91 537 37 69) o comunicacion@fbbva.es o consultar en la web www.fbbva.es